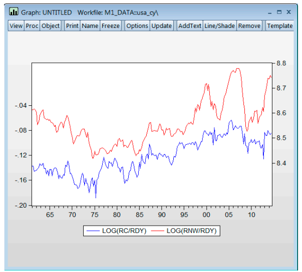
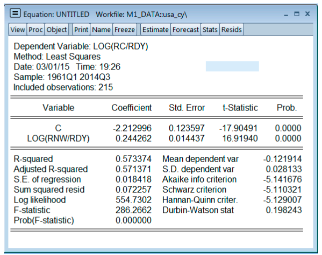
Case Study: Predicting U.S. Saving Behavior after the 2008 Financial Crisis
(proposed solution)

1. Data on U.S. consumption, income, and saving for 1947:1 – 2014:3 can be found in MF_Data.wk1, pagefile USA_CY.

2. To get a feel for the data, plot log(rc/rdy) against log(rnw/rdy). Also regress log(rc/rdy) on a constant and log(rnw/rdy) and then examine the fitted residuals, looking for evidence of any changes in the relationship between consumption, disposable income and net wealth. Is there a structural break in the relationship before 1975:4? If so, use the dummy variable called sb_1975_4 to allow for it.





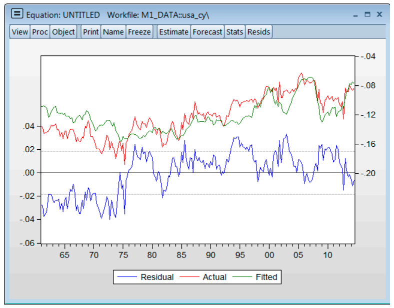
Equation: UNTITLED Workfile: M1_DATA::usa_cy\

Dependent Variable: LOG(RC/RDY)
Method: Least Squares
Date: 03/01/15 Time: 19:26
Sample: 1961Q1 2014Q3
Included observations: 215

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.212996	0.123597	-17.90491	0.0000
LOG(RNW/RDY)	0.244262	0.014437	16.91940	0.0000

R-squared	0.573374	Mean dependent var	-0.121914
Adjusted R-squared	0.571371	S.D. dependent var	0.028133
S.E. of regression	0.018418	Akaike info criterion	-5.141676
Sum squared resid	0.072257	Schwarz criterion	-5.110321
Log likelihood	554.7302	Hannan-Quinn criter.	-5.129007
F-statistic	286.2662	Durbin-Watson stat	0.198243
Prob(F-statistic)	0.000000		

Both the time plot (graph above) and the regression of log(rc/rdy) on a constant and log(rnw/rdy) (regression results shown left) suggests that there was a structural break around 1975:4 (notice the consistently negative residuals till 1975:4 (see graph below) and the wider gap between log(rc/rdy) and log(rnw/rdy) before 1976 compared to the gap after 1976). There is strong suggestive evidence that a structural break occurred around 1975:4.



3. Specify the long-run predictive equation for *real* consumption. Be sure to state your priors for the correct signs of the coefficients. It would be helpful if your model was expressed in "log" form, so that the parameter estimates are elasticities.

Consider the following long-run model for U.S. real consumption:

$$\ln rc_t = \beta_0 + \beta_1 \ln rdy_t + \beta_2 \ln rnw_t + \epsilon_t$$

where **rc** is real consumption, **rdy** is real disposable income, and **rnw** is real net worth.

- Generate real net worth: Deflate nominal net worth using the GDP deflator (**y_deflator**) as both consumption and disposable income are also deflated by y_deflator.

  genr rnw = net_worth/y_deflator

4. The sample period for model specification (i.e., specifying the preferred model for forecasting purposes) can be 1961:1 to 2007:4 (and no later than 2007:4).

- Specify the sample period using the "smpl" command.

  Smpl 1961:1 2007:4

5. Determine the statistical properties of the included variables (I(0) versus I(1)), formally test for unit roots in the data and cointegration if appropriate, and then estimate the long-run model using an appropriate estimator.

Unit root test for individual variables:

Command	**Results**
uroot log(rc)	: log(rc) has a unit root (p-value = 0.5185)
uroot log(rdy)	: log(rdy) has a unit root (p-value = 0.1246)
uroot log(rnw)	: log(rnw) has a unit root (p-value = 0.9789)

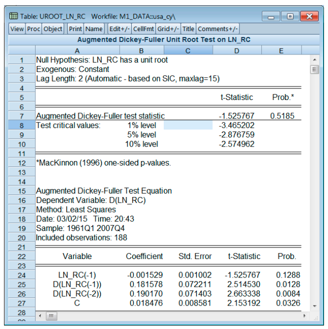


Table: UROOT_LN_RC Workfile: M1_DATA::usa_cy\

Augmented Dickey-Fuller Unit Root Test on LN_RC

Null Hypothesis: LN_RC has a unit root
Exogenous: Constant
Lag Length: 2 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.525767	0.5185
Test critical values:	1% level	-3.465202	
	5% level	-2.876759	
	10% level	-2.574962	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LN_RC)
Method: Least Squares
Date: 03/02/15 Time: 20:43
Sample: 1961Q1 2007Q4
Included observations: 188

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_RC(-1)	-0.001529	0.001002	-1.525767	0.1288
D(LN_RC(-1))	0.181578	0.072211	2.514530	0.0128
D(LN_RC(-2))	0.190170	0.071403	2.663338	0.0084
C	0.018476	0.008581	2.153192	0.0326

The null hypothesis of a unit root cannot be rejected for any of the variables using the augmented Dickey-Fuller statistic.[1] We can therefore ask whether the variables form a cointegrated system with a given number of "common trends".

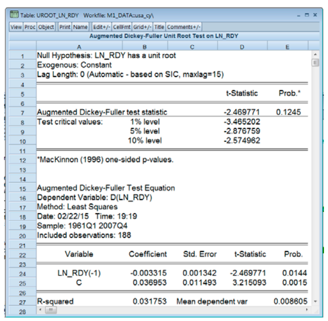


Table: UROOT_LN_RDY Workfile: M1_DATA::usa_cy\

Augmented Dickey-Fuller Unit Root Test on LN_RDY

Null Hypothesis: LN_RDY has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.469771	0.1245
Test critical values:	1% level	-3.465202	
	5% level	-2.876759	
	10% level	-2.574962	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LN_RDY)
Method: Least Squares
Date: 02/22/15 Time: 19:19
Sample: 1961Q1 2007Q4
Included observations: 188

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_RDY(-1)	-0.003315	0.001342	-2.469771	0.0144
C	0.036953	0.011493	3.215093	0.0015
R-squared	0.031753	Mean dependent var		0.008605

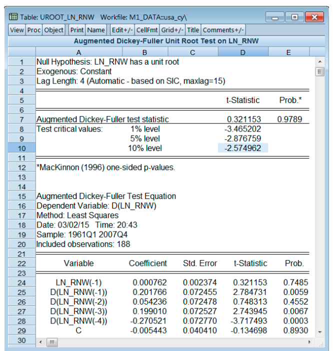


Table: UROOT_LN_RNW Workfile: M1_DATA::usa_cy\

Augmented Dickey-Fuller Unit Root Test on LN_RNW

Null Hypothesis: LN_RNW has a unit root
Exogenous: Constant
Lag Length: 4 (Automatic - based on SIC, maxlag=15)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		0.321153	0.9789
Test critical values:	1% level	-3.465202	
	5% level	-2.876759	
	10% level	-2.574962	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LN_RNW)
Method: Least Squares
Date: 03/02/15 Time: 20:43
Sample: 1961Q1 2007Q4
Included observations: 188

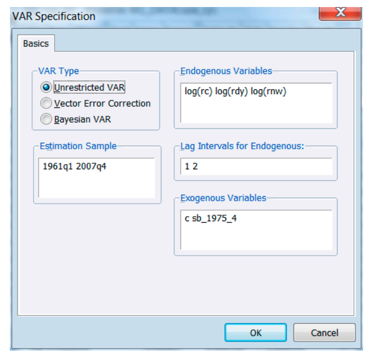
Variable	Coefficient	Std. Error	t-Statistic	Prob.
LN_RNW(-1)	0.000762	0.002374	0.321153	0.7485
D(LN_RNW(-1))	0.201766	0.072455	2.784731	0.0059
D(LN_RNW(-2))	0.054236	0.072478	0.748313	0.4552
D(LN_RNW(-3))	0.199010	0.072527	2.743945	0.0067
D(LN_RNW(-4))	-0.270521	0.072770	-3.717493	0.0003
C	-0.005443	0.040410	-0.134698	0.8930

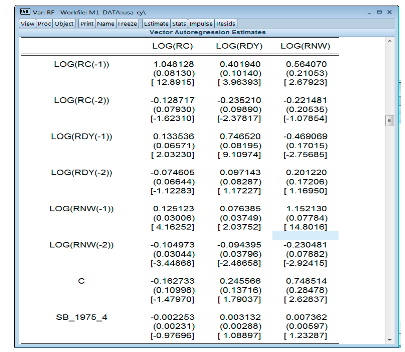
[1] The ADF results are confirmed (i.e., in favor of the unit root hypothesis) if one uses the KPSS statistic (stationary null, assuming a constant and deterministic time trend), or the same tests allowing for a structural break in either the intercept or the intercept and trend.

Cointegration test:

Given that all variables ($\ln rc_t$, $\ln rdy_t$, and $\ln rnw_t$) are consistent with the I(1) hypothesis and possibly cointegrated, we can use Johansen's trace/max eigenvalue statistic to test for the existence of a common trend (i.e., a long-run relationship).

First, estimate the VAR model of $\ln rc_t$, $\ln rdy_t$, and $\ln rnw_t$, initially using an ad-hoc number of lags, say 2. (Write "**VAR**" in the command window and press "Enter"). We also include the dummy variable sb_1975_4 to allow for the possibility of a structural break prior to 1976. The VAR object is called "**rf**" in the workfile.




Var: RF Workfile: M1_DATA::usa_cy\

Vector Autoregression Estimates

	LOG(RC)	LOG(RDY)	LOG(RNW)
LOG(RC(-1))	1.048128	0.401940	0.564070
	(0.08130)	(0.10140)	(0.21053)
	[12.8915]	[3.96393]	[2.67923]
LOG(RC(-2))	-0.128717	-0.235210	-0.221481
	(0.07930)	(0.09890)	(0.20535)
	[-1.62310]	[-2.37817]	[-1.07854]
LOG(RDY(-1))	0.133536	0.746520	-0.469069
	(0.06571)	(0.08195)	(0.17015)
	[2.03230]	[9.10974]	[-2.75685]
LOG(RDY(-2))	-0.074605	0.097143	0.201220
	(0.06644)	(0.08287)	(0.17206)
	[-1.12283]	[1.17227]	[1.16950]
LOG(RNW(-1))	0.125123	0.076385	1.152130
	(0.03006)	(0.03749)	(0.07784)
	[4.16252]	[2.03752]	[14.8016]
LOG(RNW(-2))	-0.104973	-0.094395	-0.230481
	(0.03044)	(0.03796)	(0.07882)
	[-3.44868]	[-2.48658]	[-2.92415]
C	-0.162733	0.245566	0.748514
	(0.10998)	(0.13716)	(0.28478)
	[-1.47970]	[1.79037]	[2.62837]
SB_1975_4	-0.002253	0.003132	0.007362
	(0.00231)	(0.00288)	(0.00597)
	[-0.97696]	[1.08897]	[1.23287]

Also determine the optimal number of lags using the SIC and AIC criteria. (**View>Lag Structure> Lag length criteria**)

The SC statistic suggests one lag for the optimal number of lags, whereas the AIC, HQ, and FPE statistics suggest 2. We choose the VAR model with two lags to allow for the possibility

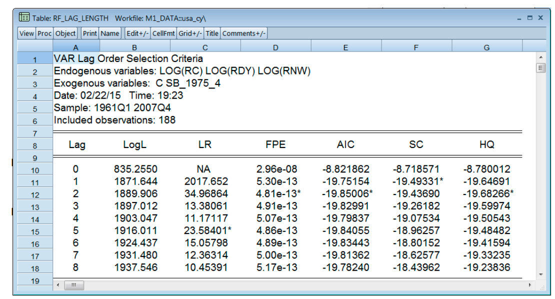


Table: RF_LAG_LENGTH Workfile: M1_DATA::usa_cy\

VAR Lag Order Selection Criteria
Endogenous variables: LOG(RC) LOG(RDY) LOG(RNW)
Exogenous variables: C SB_1975_4
Date: 02/22/15 Time: 19:23
Sample: 1961Q1 2007Q4
Included observations: 188

Lag	LogL	LR	FPE	AIC	SC	HQ
0	835.2550	NA	2.96e-08	-8.821862	-8.718571	-8.780012
1	1871.644	2017.652	5.30e-13	-19.75154	-19.49331*	-19.64691
2	1889.906	34.96864	4.81e-13*	-19.85006*	-19.43690	-19.68266*
3	1897.012	13.38061	4.91e-13	-19.82991	-19.26182	-19.59974
4	1903.047	11.17117	5.07e-13	-19.79837	-19.07534	-19.50543
5	1916.011	23.58401*	4.86e-13	-19.84055	-18.96257	-19.48482
6	1924.437	15.05798	4.89e-13	-19.83443	-18.80152	-19.41594
7	1931.480	12.36314	5.00e-13	-19.81362	-18.62577	-19.33235
8	1937.546	10.45391	5.17e-13	-19.78240	-18.43962	-19.23836

The next step is to test the residuals for autocorrelation, heteroskedasticity and normality. For the LM-test for autocorrelation (**View>Residual Tests>Autocorreation LM Test**) select lag order 12, so that we can check the null of no autocorrelation for lag orders up to 2, 4, 8 and 12 - reasonable lag orders given that we work with quarterly data.

VAR Residual Serial Correlation LM T...
Null Hypothesis: no serial correlation at ...
Date: 04/29/15 Time: 08:46
Sample: 1961Q1 2007Q4
Included observations: 188

Lags	LM-Stat	Prob
1	13.60887	0.1369
2	8.080797	0.5260
3	21.40155	0.0110
4	24.25051	0.0039
5	10.65281	0.3003
6	6.404718	0.6988
7	6.217350	0.7180
8	14.69451	0.0997
9	3.494151	0.9415
10	8.069861	0.5271
11	11.93950	0.2167
12	7.993124	0.5348

Probs from chi-square with 9 df.

The LM test suggests that we cannot reject the null of no autocorrelation for lag orders up to 2, 8 and 12. The test, however, rejects no autocorrelation for lag order up to 4. In such a case, you may want to either proceed with the model as is, relying on the test results for higher lag orders (8 and 12), or think about possible ways of how to amend the model and deal with the signs of possible autocorrelation at lag orders 3 and 4. We choose to proceed with the model as is, i.e. a VAR(2).

There are two possible test specifications to check residual heteroskedasticity: with (**View>Residual Tests>White Heteroskedasticity (No Cross Terms)**) or without cross-terms (**View>Residual Tests>White Heteroskedasticity (With Cross Terms)**) in the test model. The results are presented below.

VAR Residual Heteroskedasticity Tests: No Cross Terms (only levels and squares)
Date: 04/29/15 Time: 08:51
Sample: 1961Q1 2007Q4
Included observations: 188

Joint test:

Chi-sq	df	Prob.
101.0340	78	0.0409

Individual components:

Dependent	R-squared	F(13,174)	Prob.	Chi-sq(13)	Prob.
res1*res1	0.116413	1.763434	0.0523	21.88569	0.0572
res2*res2	0.121184	1.845667	0.0396	22.78260	0.0444
res3*res3	0.143907	2.249920	0.0094	27.05452	0.0122
res2*res1	0.075922	1.099678	0.3622	14.27336	0.3549
res3*res1	0.093777	1.385048	0.1706	17.63000	0.1721
res3*res2	0.086720	1.270927	0.2346	16.30334	0.2331

VAR Residual Heteroskedasticity Tests: Includes Cross Terms
Date: 04/29/15 Time: 08:53
Sample: 1961Q1 2007Q4
Included observations: 188

Joint test:

Chi-sq	df	Prob.
174.3249	138	0.0198

Individual components:

Dependent	R-squared	F(23,164)	Prob.	Chi-sq(23)	Prob.
res1*res1	0.189984	1.672400	0.0349	35.71704	0.0441
res2*res2	0.258297	2.483157	0.0005	48.55974	0.0014
res3*res3	0.184410	1.612237	0.0463	34.66910	0.0561
res2*res1	0.131433	1.078993	0.3737	24.70949	0.3654
res3*res1	0.140953	1.169968	0.2791	26.49919	0.2779
res3*res2	0.126293	1.030691	0.4304	23.74303	0.4182

In both cases the null of homoskedasticity is rejected at 5%. Heteroskedastic errors do not affect model forecasting performance in terms of mean forecast, which is what we use our model for at a later stage of the assignment. However, if you decide to account for heteroskedasticity, you would need to reformulate your model as a system and use multivariate GARCH specification to model the variance-covariance matrix (this however goes beyond the scope of this course).

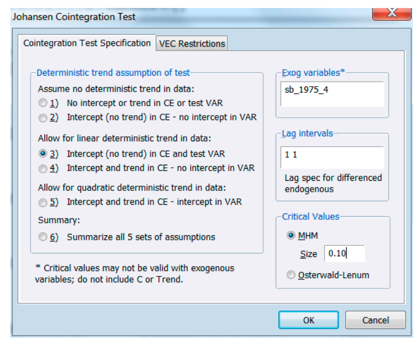
VAR Residual Normality Tests
Orthogonalization: Cholesky (Lutkepohl)
Null Hypothesis: residuals are multivariate normal
Date: 04/29/15 Time: 08:56
Sample: 1961Q1 2007Q4
Included observations: 188

Component	Skewness	Chi-sq	df	Prob.
1	-0.606544	11.52740	1	0.0007
2	0.493714	7.637607	1	0.0057
3	-0.459070	6.603342	1	0.0102
Joint		25.76835	3	0.0000

Component	Kurtosis	Chi-sq	df	Prob.
1	5.040982	32.63059	1	0.0000
2	5.273215	40.47879	1	0.0000
3	3.898877	6.329178	1	0.0119
Joint		79.43856	3	0.0000

Component	Jarque-Bera	df	Prob.
1	44.15799	2	0.0000
2	48.11640	2	0.0000
3	12.93252	2	0.0016
Joint	105.2069	6	0.0000

The Jarque-Bera test for normality (**View>Residual Tests>Normality Test>Cholesky of Covariance**) rejects the null hypothesis of joint normal distribution of errors. Note that the non-normality finding may be related to the presence of heteroskedasticity. The latter may explain excess kurtosis in residual distribution. This however, does not seem to be the only possible cause: also, residuals from the preferred model seem to be skewed.



Then, we conduct the Johansen's test with one lag. (**View>Cointegration Test**) As explained in the lectures, please ensure that the lag interval for the **differenced** endogenous variables is 1 smaller than the number of lags used in by moving to the corresponding VECM form (as explained in Module 6). Note that we have also included the structural break (sb_1975_4) variable as an exogenous variable.

The results suggest that we can reject the null hypothesis of **no cointegration** (or zero cointegrating vectors) using a 6 percent significance level. Moreover, we cannot reject the null hypotheses of at most 1 cointegrating vectors versus the alternative of 2. **Therefore, we assume that there exists one (and only one) cointegrating vector.**

Johansen Cointegration Test

Date: 02/22/15 Time: 19:39
Sample: 1961Q1 2007Q4
Included observations: 188
Trend assumption: Linear deterministic trend
Series: LOG(RC) LOG(RDY) LOG(RNW)
Exogenous series: SB_1975_4
Warning: Critical values assume no exogenous series
Lags interval (in first differences): 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.1 Critical Value	Prob.**
None *	0.121551	29.37115	27.06695	0.0559
At most 1	0.023458	5.006875	13.42878	0.8081
At most 2	0.002890	0.544155	2.705545	0.4607

Trace test indicates 1 cointegrating eqn(s) at the 0.1 level
* denotes rejection of the hypothesis at the 0.1 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.1 Critical Value	Prob.**
None *	0.121551	24.36428	18.89282	0.0169
At most 1	0.023458	4.462720	12.29652	0.8076
At most 2	0.002890	0.544155	2.705545	0.4607

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.1 level
* denotes rejection of the hypothesis at the 0.1 level
**MacKinnon-Haug-Michelis (1999) p-values

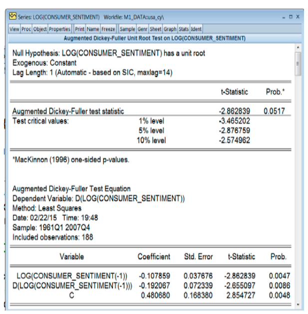
6. Starting from the "long-run" specification, construct an ECM equation (e.g., an error-correction model) that can explain the actual behavior of consumption during the sample period. Justify the ECM specification using standard regression diagnostics, including AIC/SIC, robustness of the parameter estimates, and residual diagnostics (actual versus fitted).

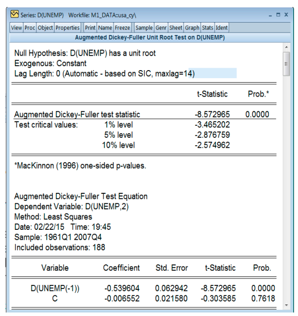
In developing the ECM model, feel free to include additional exogenous variables in the VECM to explain short-run dynamics of consumption. Here we use the unemployment rate and the consumer confidence index as exogenous variables. These variables are proxies for changes in the level of income uncertainty facing the household sector.

Let's first conduct a unit root test for (1) the level of consumer sentiment and (2) the change in the unemployment rate and to check that they are separately consistent with the stationarity assumption. The required EViews commands are:

Command	**Results**
uroot d(unemp)	d(unemp) is stationary (p-value = 0.0000)
uroot log(consumer_sentiment)	log(consumer_sentiment) is stationary (p-value= 0.0517)

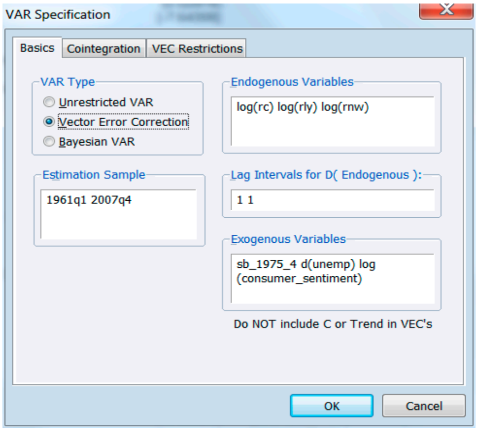
Both *d(unemp)* and *log(consumer_sentiment)* are consistent with the stationarity hypothesis.[2]

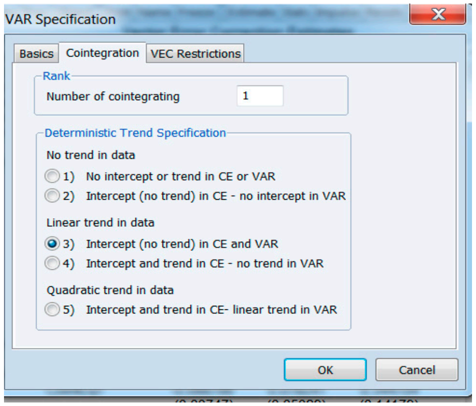




[2] Technically, log(consumer_sentitiment) is consistent with the unit root hypothesis using a 5 percent level of significance. However, the estimated first-order autoregressive coefficient is quite far from unity (0.89), raising doubt about the power of the augmented Dickey-Fuller unit root test in this context.

The next step is to estimate the corresponding ECM model using the VECM procedure. Type "**var**" in the command window and press the "enter" key. Then, specify the form of the VEC model in the resulting pop-up window, allowing for the three exogenous variables (sb_1975_4, d(unemp) and log(consumer_sentiment))[3].

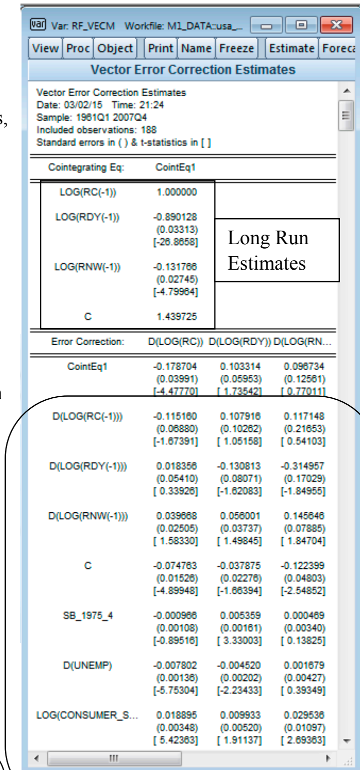




[3] The cointegration test is even more in favor of one cointegrating vector if we add d(unemp) and log(consumer_sentiment) to the exogenous variable list. Currently, it contains only sb_1975_4.

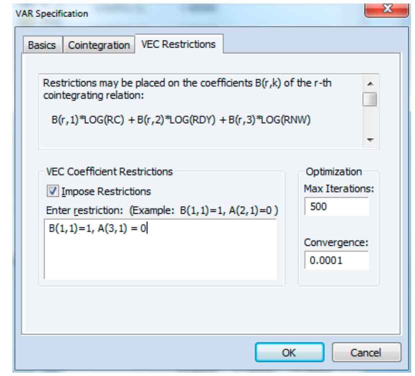
Again, please ensure that the lag intervals for the differenced **endogenous** variables are 1 (given we estimated the VAR with 2 lags in levels). Also, you will need to click on the **cointegration** tab to specify the number of cointegrating vectors, which in this case is 1. Also indicate the form of the deterministic part in the ECM regression, typically option 3.

The coefficients in the long-run equation (see right image) are both statistically significantly different from zero (see the right figure), and most importantly consistent with our economic priors (i.e. income and wealth have positive coefficients on consumption; see module 6). Also, the coefficient on the ECM (**cointeq1**) term is negative and significant with respect to real consumption, and significant and positive with respect to disposable income, both using a 5% (one-sided) level of significance. In contrast, the insignificance of the ECM coefficient in the real wealth equation suggests that real wealth is weakly exogenous with respect to consumption. Some would argue that these results are reasonable since both consumption and income are endogenous (flow) variables and dependent on each other, while real wealth (a stock) takes time to react to disequilibria in consumption relative to its long-run path.



Short-run parameter estimates (standard VAR if we ignore CointEq1)

To illustrate how to impose restrictions on the VECM, we now test for the weak exogeneity of log(rnw), by imposing "A(3,1) = 0 on the fitted model. Recall that we can work with the "VEC restrictions" tab (shown above right) to do this efficiently.

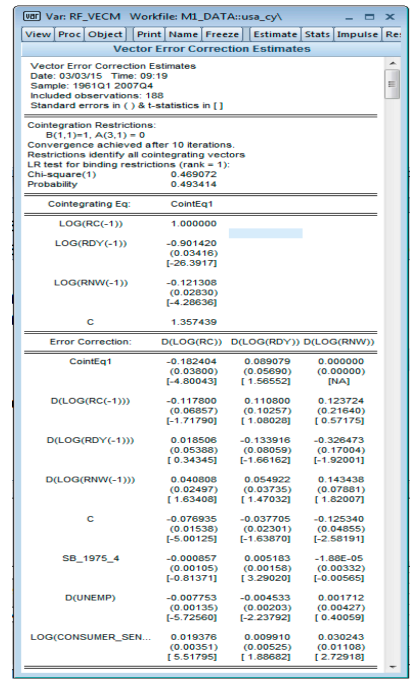


Note that to impose restrictions on the VECM you need to indicate – at the very least - which variable is to be placed on the left-hand side of each long-run co-integrating equation (of course, there may be more than one). This is the purpose of "B(1,1)=1" in the restriction text box above (see right image). The statement instructs EViews to treat **rc** as the dependent variable in the long-run equation.

The results of imposing the restriction are shown in the right-hand screen shot. The p-value of the test statistic for weak exogeneity is 0.4934. Thus we can accept the null. That is log(rnw) can be treated as weakly exogenous.

We will therefore continue use this restriction in our baseline specification of the model.

Algebraically, the preferred model has the following form (the common error correction term is highlighted in bold):



Consumption:

d(log(rc)) = -0.1824 * (**log(rc(-1)) - 0.9014 * log(rdy(-1)) - 0.1213 * log(rnw(-1)) - 1.3574**) - 0.1178 * d(log(rc(-1))) + 0.0185 * d(log(rdy(-1))) + 0.0408 * d(log(rnw(-1))) - 0.0769 - 0.00086* sb_1975_4 - 0.00775 * d(unemp) + 0.0194 * log(consumer_sentiment)

Real Disposable Income:

d(log(rdy)) = 0.089079 * (**log(rc(-1)) - 0.9014 * log(rdy(-1)) - 0.1213 * log(rnw(-1)) - 1.3574**) + 0.1108 * d(log(rc(-1))) - 0.1339 * d(log(rdy(-1))) + 0.0592 * d(log(rnw(-1))) - 0.0377 + 0.0052 * sb_1975_4 -0.00453 * d(unemp) + 0.0099 * log(consumer_sentiment)

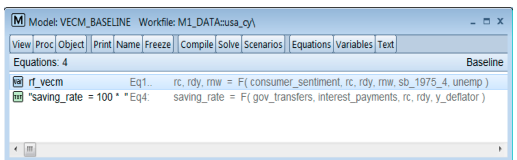
Real Net Worth:

d(log(rnw)) = 0.0 * (**log(rc(-1)) - 0.9014 * log(rdy(-1)) - 0.1213 * log(rnw(-1)) - 1.3574**) + 0.1237* d(log(rc(-1))) - 0.3265 * d(log(rdy(-1))) + 0.1435 * d(log(rnw(-1))) -0.1253 - 1.88E-05 * sb_1975_4 + 0.00171 * d(unemp) + 0.0302 * log(consumer_sentiment)

Notice that the coefficient on the ECM term in the real net worth equation has been constrained to zero as required, reflecting the results of the weak exoeneity test.

7. Use your preferred model and the EViews' solver to separately produce dynamic, static, and "fit" forecasts for the *U.S. saving rate*[4] for 2008:1-2014:3 (depending on availability of data). Given your findings, discuss whether the recent upswing in the U.S. saving rate is permanent or transitory.

You can generate the forecasts for the U.S. saving rate for 2008:1 - 2014:3 using EView's simulator.



[4] The saving ratio (percentage points) can be derived from total nominal consumption and nominal household disposable income using the following formula:

saving rate=100*(rdy - (rc + ((gov_transfers + interest_payments/y_deflator))))/rdy

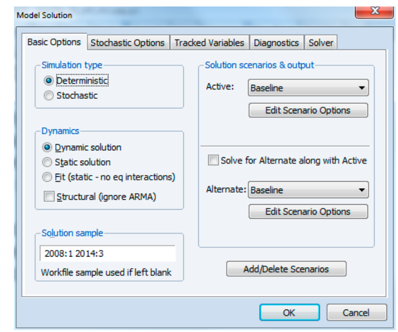
First create a model object using (**Proc> Make Model**) from the baseline VECM model.

In order to produce the forecasts for the saving rate from the estimated consumption function, we need to add the relevant accounting identity for the saving rate. To do so, open the model object, and then press **Right click > Insert > "Edit Equation"** to add the following identity

saving_rate=100*(rdy - (rc + ((gov_transfers + interest_payments/y_deflator))))/rdy

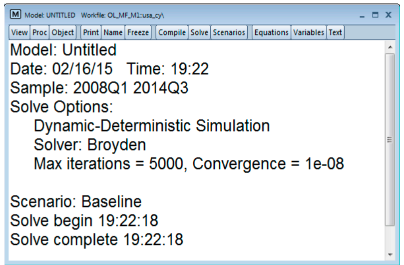
This definition is adapted from the one used in the U.S. National Income and Product Accounts (NIPA), Table 2.1.

Note that the definition must be in terms of the same variables being predicted by the model (i.e., real consumption (rc) and real disposable income(rdy)). Otherwise the forecasted values of the real variables will not be used to predict the saving ratio during the forecast period. (A common mistake is to use the definition of the saving rate using nominal quantities.)



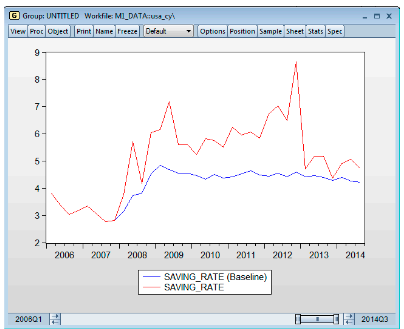
Calculate the deterministic (or point) baseline forecast series for 2008:1 – 2014:3 using "solve" command. Choose "dynamic solution".

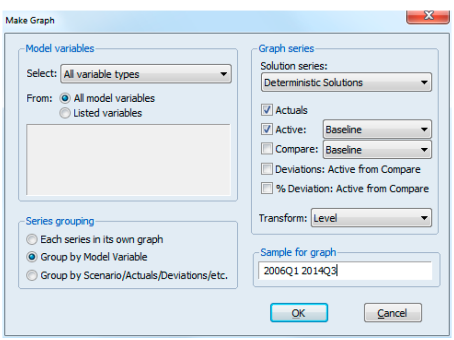
The baseline forecasts for all the endogenous variables in the model will be saved to your workfile with "_0" extensions. Thus, the forecast for real consumption can be found in the series called rc_0. The baseline forecast for the saving rate will be in the series called **saving_rate_0**.



After the solution is created (see left image), you can plot the baseline (forecasts) with the actual series for real consumption and the saving rate using

line rc rc_0, and **line saving_rate saving_rate_0**).

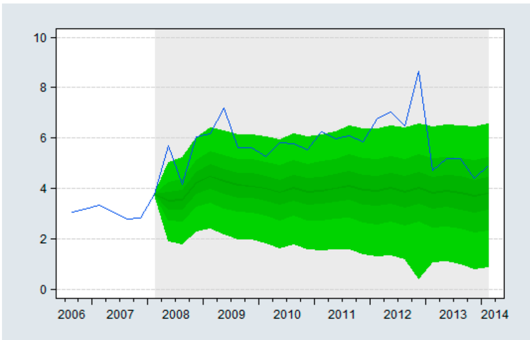




Alternatively, you can easily plot all the variables against their actual outcomes using **Proc>Make Graph** after the model is solved.

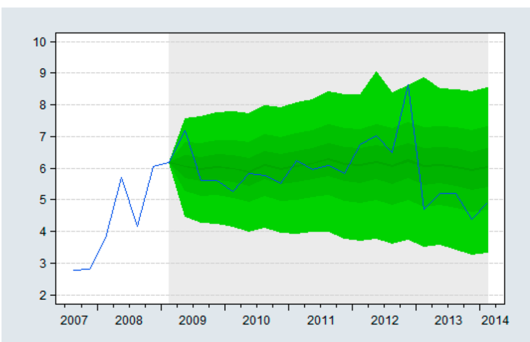
Not surprisingly, the dynamic forecasts (green lines) are over predicting actual real disposable income, real net wealth and therefore real consumption expenditures (blue lines). This reflects the I(1) nature of the variables involved, the limited serial correlation in their innovation sequences, and the fact that – prior to 2008 – both variables were trending upwards. Because of their I(1) nature, a dynamic forecast will effectively use the last known value of these variables to generate out-of-sample forecasts. Consequently, the forecasting model does not do good job of predicting the impact of the global financial crisis on real consumption, and its predictions for the saving ratio appear to be systematically below the actual saving outcome.

That said, the previous assessment considers only a single, dynamic simulation using the baseline model. A fairer assessment of the model can be made using EViews' stochastic simulator to produce a fan chart. Using the code in m4_fanchart.prg (see module 4), the resulting fan chart (or 90 percent confidence band) for the saving rate is:



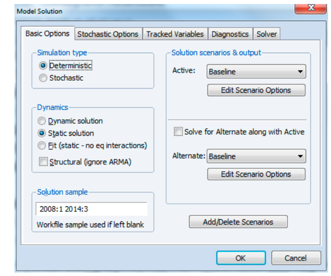
Despite the baseline model's under prediction of real net worth, the fan chart above encompasses the actual outcomes of the saving rate for much of 2008:1-2014:3, suggesting that the underlying model is useful for predicting the outcome of the saving rate.

While there is obvious (downward) bias in the model's predictions, it should be noted that this partly a reflection of the initial starting point of the simulation experiment. If one begins the simulations in 2009:1, the resulting fan chart is:



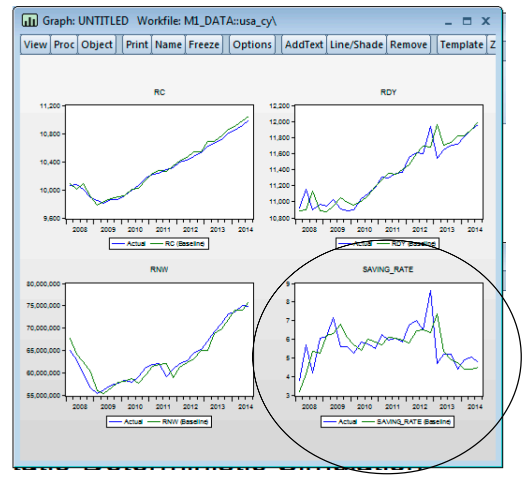
Consequently, if the estimated model uses the actual outcomes of disposable income and real wealth for 2008, even a dynamic forecast can yield a confidence band that successfully encompasses the actual behavior of the saving rate.

This result can be confirmed by solving the model using the "**Static**" method, which is equivalent to one-period ahead forecasting. In other words, the simulation uses the last known values of the RHS endogenous variables (lagged consumption, lagged income, and lagged real net worth) to calculate the baseline forecast. We would naturally expect this forecast to be more accurate than the previous dynamic forecast, where the forecast errors of the endogenous variables are allowed to accumulate throughout the forecast horizon.



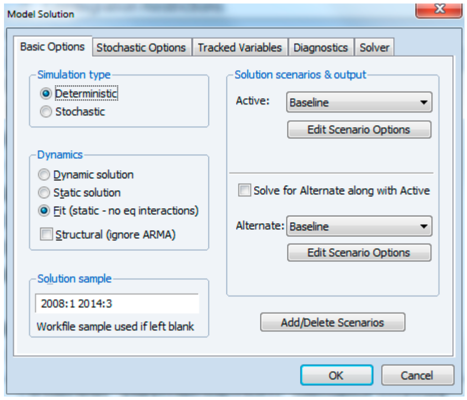
The results are shown in the adjacent figure (right). The consumption forecast (rc) from the static method is indeed better than that obtained using a dynamic simulation.

Given that we are conducting an out-of-sample forecast, the superior forecast of the static solution implies that the forecasting error in the consumption equation is largely from the errors in forecasting disposable income and net worth, not from changes in the estimated parameters.

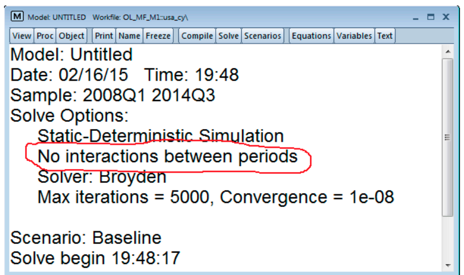


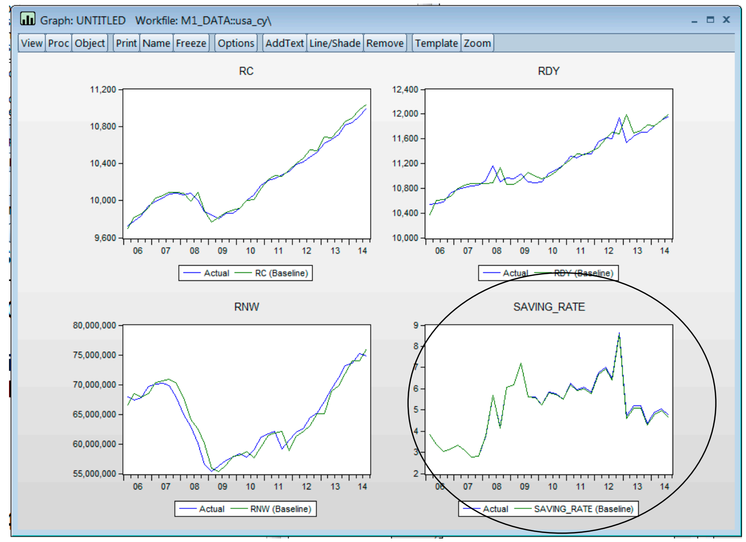
This experiment provides suggestive evidence that there is no structural change in the consumption function/saving behavior. The increase in the saving ratio can be predicted rather well using the (lagged, actual) values of real income and real net worth.

This hypothesis can be taken to an extreme by calculating the so called "single equation forecast" of the system. This is the forecast that ignores all the interactions between the endogenous variables in the model. In this mode, the simulator is in effect using the actual values of income and net worth to produce the forecast of consumption, and hence, the saving ratio. Moreover, the dynamic interdependencies in the system are "turned-off" so that there are no equation interactions between periods.



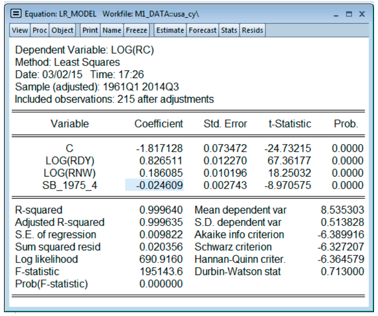
A comparison of the "fit" forecast against the actual outcome of the saving ratio will reveal whether the model – with parameter estimates derived using data prior to the crisis – can forecast the out-of-sample saving ratio. It is essentially the Chow Forecast Test for structural stability, using parameter estimates derived from a system estimator (i.e., VECM).





Notice that the ECM model, in single equation mode and using parameter estimates using data for a period that did not encompass the great financial crisis, can predict the actual behavior of the saving rate rather well. This result provides further evidence that the major driving forces behind the increase in the saving ratio during the great recession were the negative shocks to income and net worth, not a change in the underlying saving behavior (specifically, the response of real consumption to income and net worth) of the American household. It follows that we should expect the saving ratio to revert to its norm (i.e., lower) level once real income (and wealth) levels recover to pre-crisis levels.

Lastly, let's use the basic long-run equation involving log(rc), log(rdy) and log(rnw) to perform a simple out-of-sample forecasting experiment. The result of estimating this regression from 1961Q4 to 2014Q3 using ordinary least squares (other estimators that allow for the I(1) nature of the variables give similar results) is:

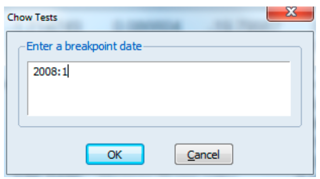

Equation: LR_MODEL Workfile: M1_DATA::usa_cy\

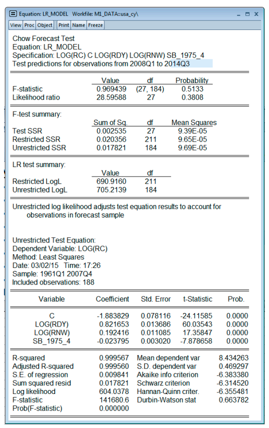
Dependent Variable: LOG(RC)
Method: Least Squares
Date: 03/02/15 Time: 17:26
Sample (adjusted): 1961Q1 2014Q3
Included observations: 215 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.817128	0.073472	-24.73215	0.0000
LOG(RDY)	0.826511	0.012270	67.36177	0.0000
LOG(RNW)	0.186085	0.010196	18.25032	0.0000
SB_1975_4	-0.024609	0.002743	-8.970575	0.0000

R-squared	0.999640	Mean dependent var	8.535303
Adjusted R-squared	0.999635	S.D. dependent var	0.513828
S.E. of regression	0.009822	Akaike info criterion	-6.389916
Sum squared resid	0.020356	Schwarz criterion	-6.327207
Log likelihood	690.9160	Hannan-Quinn criter.	-6.364579
F-statistic	195143.6	Durbin-Watson stat	0.713000
Prob(F-statistic)	0.000000		

Now conduct a "Chow Forecast Test", which calculates an F-statistic that assesses whether the forecasts from the within sample regression (above) can predict out-of-sample results (in this case, real consumption for 2008Q1 – 2014Q3). **View> Stability Test> Chow Forecast Test.** Set the breakpoint date at 2008Q1.




Equation: LR_MODEL Workfile: M1_DATA::usa_cy\

Chow Forecast Test
Equation: LR_MODEL
Specification: LOG(RC) C LOG(RDY) LOG(RNW) SB_1975_4
Test predictions for observations from 2008Q1 to 2014Q3

	Value	df	Probability
F-statistic	0.969439	(27, 184)	0.5133
Likelihood ratio	28.59588	27	0.3808

F-test summary:

	Sum of Sq.	df	Mean Squares
Test SSR	0.002535	27	9.39E-05
Restricted SSR	0.020356	211	9.65E-05
Unrestricted SSR	0.017821	184	9.69E-05

LR test summary:

	Value	df
Restricted LogL	690.9160	211
Unrestricted LogL	705.2139	184

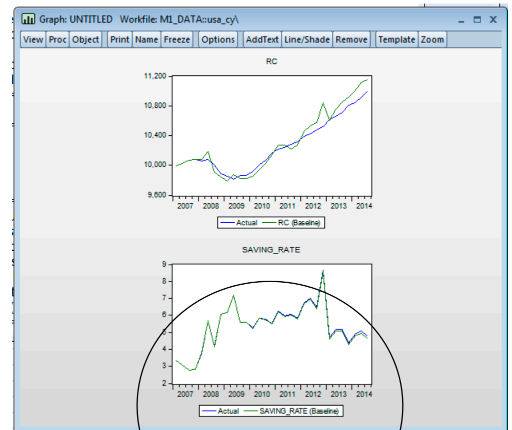
Unrestricted log likelihood adjusts test equation results to account for observations in forecast sample

Unrestricted Test Equation:
Dependent Variable: LOG(RC)
Method: Least Squares
Date: 03/02/15 Time: 17:26
Sample: 1961Q1 2007Q4
Included observations: 188

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.883829	0.078116	-24.11585	0.0000
LOG(RDY)	0.821653	0.013686	60.03543	0.0000
LOG(RNW)	0.192416	0.011085	17.35847	0.0000
SB_1975_4	-0.023795	0.003020	-7.878658	0.0000

R-squared	0.999567	Mean dependent var	8.434263
Adjusted R-squared	0.999560	S.D. dependent var	0.469297
S.E. of regression	0.009841	Akaike info criterion	-6.383380
Sum squared resid	0.017821	Schwarz criterion	-6.314520
Log likelihood	604.0378	Hannan-Quinn criter.	-6.355481
F-statistic	141680.6	Durbin-Watson stat	0.663782
Prob(F-statistic)	0.000000		

The small value of the F-statistic suggests that there is no structural break in the long-run relationship. This can be confirmed by using this regression to predict the out-of-sample saving ratio. The results of doing so are shown on the next page.



The long-run model without any short-run dynamics is able to produce acceptable out-of-sample forecasts of real consumption and the household saving ratio, providing further evidence that there has **not** been a structural change in U.S. household consumption habits because of the global financial crisis.

Broad Conclusions:

(a) The underlying relationship between real consumption, real disposable income and real net worth did not change because of the onset of the financial crisis.

(b) The observed movements in the saving rate after 2008:1 appear to be the result of changes in disposable income and real net worth, the latter in particular being adversely affected by the financial crisis.

(c) Given that the coefficient estimates have not changed significantly, one can expect the saving rate to return to its pre-2008 level once the lingering effects of the financial crisis abate (that is, real income and real net worth return to their 2007 levels respectively). Specifically, the ratio of real net work to real disposable income at the end of 2007 was around 6.5; at the end of the sample (2014Q3), this ratio was around 6.2.

(d) Given that the consumption function did not change permanently, actions taken by the U.S. government (one being to raise government expenditure) to buffer the negative net worth shock are justifiable.