Macroeconometrics Problem Set 2: Explanatory Notes

Corti Pietro (3061891), Reichlin Vito (3081010)

14 November 2022

Point 1

a) To answer this question we firstly compute the OLS estimator and its distribution; we use a process which is very similar to the one we applied in the first problem set. In particular we firstly estimate a Random Walk process (RW) process with drift, using the command "filter" for a unit root process (so $\rho = 1$) then, thanks to a for loop we estimate the OLS coefficient, and we obtain the following distribution computing a Monte Carlo (MC) experiment:

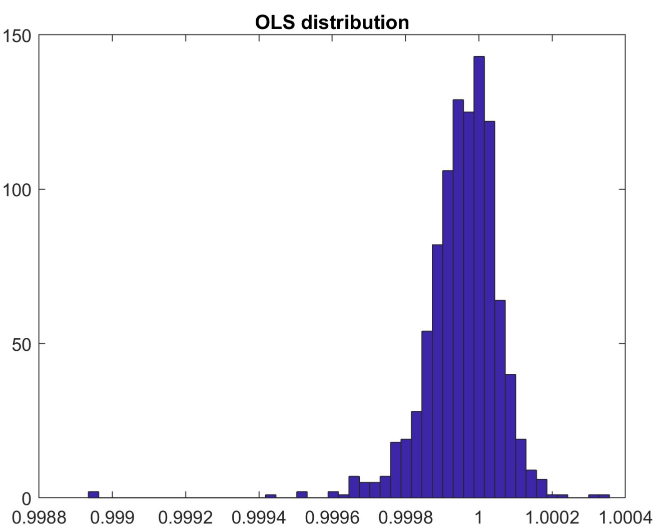

Figure 1: OLS distribution for the RW with drift

as we can see, its distribution is very similar to a Normal one, even if we should have a unit root process. This is in line with what is presented by Hamilton (1994) in chapter 17 (Case 3).

We can also have a look at the distribution of the t-stat, presented in the following figure:

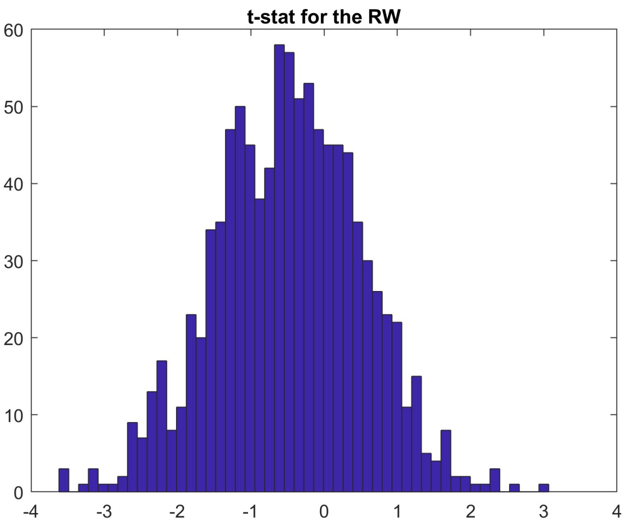

Figure 2: t-stat distribution for the RW with drift

also in this case we notice that its distribution is very similar to a Normal one. We can observe this peculiar situation because two fundamental requisites are specified:

- in the DGP we have a drift α
- in the regression estimated we also have the same constant α

Intuitively this happens because in the regression the lag regressor is asymptotically dominated by the time trend (here the constant term α); indeed, repeating many time the simulation via a MC simulation the constant term in the DGP will "go to infinity" faster by the lag term with coefficient equal to one, and in large samples this will "cover" its role in the regression.

b) Looking at the picture above we notice that the mean of the t-stat is very near to zero, or at least between -1 and 0; this fact creates a situation in which we reject very few times the null hypothesis, because basically zero times the t-stat is bigger in absolute vale than 1.96. Of course this happens because the mean value for the OLS is very near to one (see Figure 1), and so the value for this statistics is very low; as a consequence we have a test that will reject the null virtually no times, as we have remarked before.

c) If now we add a time trend component to both the DGP and the regression to estimate, we obtain the following equations (respectively for the DGP and the regression):

$$y_t = \alpha + \delta t + y_{t-1} + \epsilon_t$$

$$y_t = \alpha + \delta t + \rho y_{t-1} + \epsilon_t$$

The process that we used to estimate the OLS and the t-stat is exactly the same that we used in the previous poiints (see the matlab file); we obtain the following distribution of the t-stat:

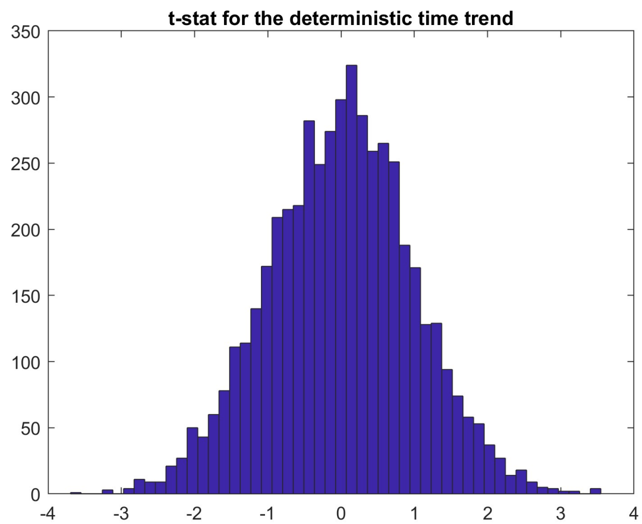

Figure 3: t-stat distribution with constant and time trend

as we can see, its distribution resembles the Normal one, with zero mean. This result is coherent with what is presented by Sims, Stock and Watson (1990). Moreover the fact that we reject the null with only a probability of 0.05 gives additional strengthen to our hypothesis.

Point 2

To answer this point we take into account what is presented by Enders Edition 4, pag. 195-199. Firstly we present the situation of the spurious regression; let's consider a regression like the following one:

$$y_t = \alpha + \beta z_t + \epsilon_t$$

where:

$$y_t = y_{t-1} + \epsilon_t^y$$

$$z_t = z_{t-1} + \epsilon_t^z$$

with both last error terms being WN and independent from each other.
It turns out that this is a case of spurious regression: the t-stat for a normality test to reject $\beta = 0$ will make no sense, as well as the index R^2. Following the same approach of Enders we consider 4 different cases to deal with this peculiar situation:

Case 1: both y_t and z_t are stationary

To study this case we compute always a MC simulation to obtain the empirical distribution of the OLS and to compute also the t-stat as well as the F-test. The computational processes are basically the same of the previous points: we generate two AR(1) processes using the function "filter" with a random term ("randn") and we obtain in this way both y_t and z_t (respectively "yt" and "zt" in our Matlab file). Then, we add a constant to the matrix of the regressors and compute in the usual way the OLS with all the components needed to calculate the statistics required (product of the error terms "ee", cariance of the error term "var_rps_hat" and the variance of the OLS estimator "var_b_hat"). Coherently with Enders explanation we find values which are correct, without any need to correct the regression; this is because y_t and z_t are both stationary. Here the empirical distribution of the OLS:

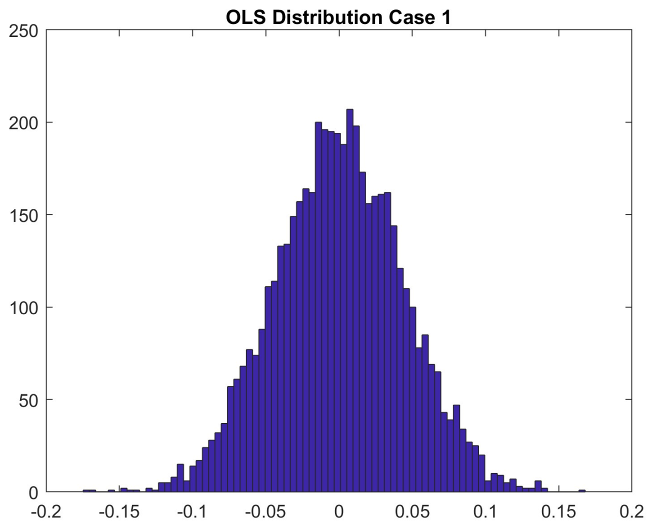

Figure 4: OLS distribution in Case 1: stationarity of both processes

which is a Normal distribution centered around 0. Also the rejection region for both the t and F tests make sense, both with a probability of approximately 5%.

We also computed the index R^2 to show that it is consistent with the "normal" case of sationarity, and indeed it gets values that are near 0.0019. It's useful to remember here how we computed it; the literature typically represents the index in this way (we consider a generic regresion $y = \beta X + \epsilon$):

$$R^2 = \frac{ESS}{TSS} = \frac{\hat{y}' M_{[I]} \hat{y}}{y' M_{[I]} y} = 1 - \frac{e'e}{y' M_{[I]} y}$$

where $\hat{y} = \hat{\beta} X$ and $M_{[I]} = I(I'I)^{-1} I'$ with I a matrix with all elements equal to unity. In our Matlab file we compute TSS in a slightly different way, exploiting the fact that the matrix $M_{[I]}$ transform all variables in deviations from their sample mean; so we express it in the following way:

$$TSS = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Of course the critical value for the F-test is not the canonical one, but the one we can use in a DF test, and it's equal to 3.84.
To sum up, in this first case no adjustment was needed since both y_t and z_t were stationary processes. As a consequence the OLS estimator could be used without nay issue.

Case 2: y_t and z_t are integrated of different order

To present this case we consider a situation in which the order of integration is $I(0)$ and $I(1)$, just for a matter of simplicity.
To express this situation in Matlab we simply consider an AR(1) process for y_t (which has of course $I(0)$) and a RW process for z_t (which is $I(1)$).
The process we used in our file to compute the OLS, the statistics and the R^2 is the same of the previous point (always a MC simulation); however now results are different: the t-test and the F-test reject the null hypothesis with a probability approximately equal to 50% (see the Matlab file), which of course doesn't make any sense since these are "useless" tests. This happens exactly because the two series are integrated of different orders. Moreover also the R^2 is very high, and so it is not very indicative, always because the two series have nothing in common. The OLS distribution is the following:

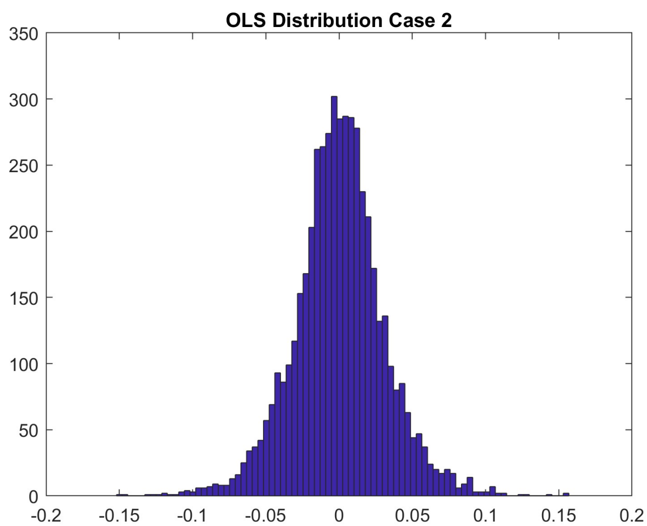

Figure 5: OLS distribution in Case 2: processes integrated of different orders

which doesn't make any sense because is too centered on the mean value, and it is not a canobical Normal distribution.

Case 3: y_t and z_t are integrated of the same order

For simplicity we work with the case in which both of the sequences are $I(1)$. As indicated by Enders, this case in the one in which the regression is spurious. Computing the OLS distribution and the test by a MC simulation we easily recognize that the results don't make sense: with both tests we reject basically all the times (more than 90% of the times, see the Matlab file) and R^2 is circa equal to 0.8, a value which of course doesn't make any sense. The OLS distribution is the following:

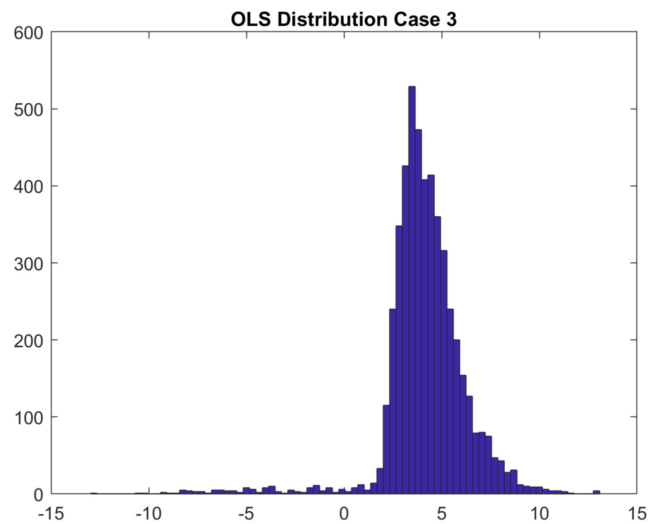

Figure 6: OLS distribution in Case 3: processes integrated of same order

It's really easy to see that this distribution has no real meaning.
This is what happens exactly in a spurious regression: with unit roots standard asymptotic inference doesn't work.

However, we can also find some ways to correct for this problems; following Enders suggestions in his book, if we take the first difference of both sequences and we make the regression using them we should be able to correct the problem. So, if we run the following regression:

$$\Delta y_t = \beta \Delta z_t + \Delta \epsilon_t$$

we obtain correct values of the statistics and of the OLS; the point is that taking the first difference allows to correct for the unit root problem.
In our Matlab file we try to make this correction in two different ways:

- we use the command "lagmatrix" and then make the difference with the leves

- we use the command "diff" which makes directly the difference with the previus lags

in the first case (the "lagmatrix approach") we use the same DGP but then we compute the lag matrices ("Yt" and "Zt" in the file), we add the constants to X, and then we run the OLS estimation of the new difference variables ("delta_X" and "delta_Y"). When we observe the values of the tests we realize that the number of times we reject H_0 decreases and it turns back to approximately 5%. The OLS distribution returns to resemble the Normal one case:

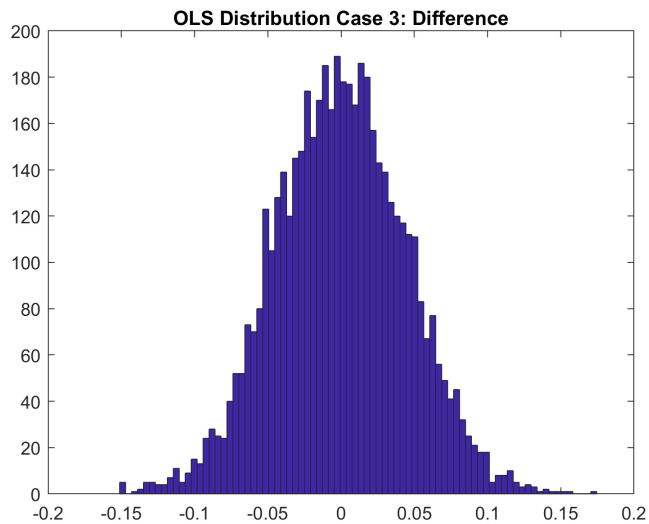

Figure 7: OLS distribution in Case 3: difference approach

If we do the second approach, the results are exactly the same; of course the way in which we settled down the variables is different: now we applied the command "diff" to the matrices to use in the regression (always adding the constant term and obtaining "diifY" for the dependent variable, and "diffX" for the independent variable). For the rest, basically nothing was changed, the command to compute the t-test anf the F-test are exactly the same, as well as the one for the R^2.

Case 4: z_t and y_t are integrated of the same order, residual is stationary

In this particular case we speak about "cointegration". A trivial case is when $\epsilon_t^y = \epsilon_t^z$, that is a case of perfect correlation. A more interesting case is the one in which:

$$y_t = \mu_t + \epsilon_t^y$$

$$z_t = \mu_t + \epsilon_t^z$$

where the error terms are WN and $\mu_t = \mu_{t-1} + \epsilon_t$, so a RW process.
In our Matlab file we do again the same procedure, changing of course the DGP and defining "mu". When we compute the test statistics, we find out that the rejection region is not indicative, given that we reject 50% of the times; moreover also the index R^2 gets values that are too big to be realistic. Lastly, here you can see that the OLS distribution has no standard distribution in the Case 4:

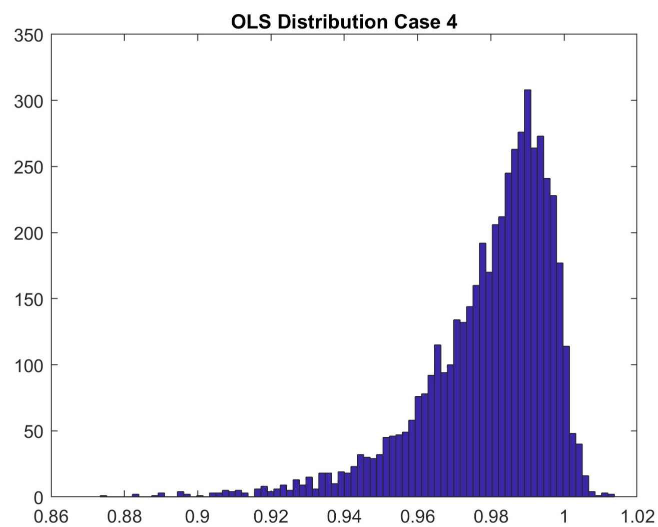

Figure 8: OLS distribution in Case 4: cointegration

Point 3

Before the solution of this point we present here the graph of all the variables used in the "Romer_Romer.xlsx" file. We have identified 3 different variables, as presented in Romer and Romer (2004): inflation, unemployment, ffr (federal fund rates) and the new monetary shock variable. Below you can see the graph:

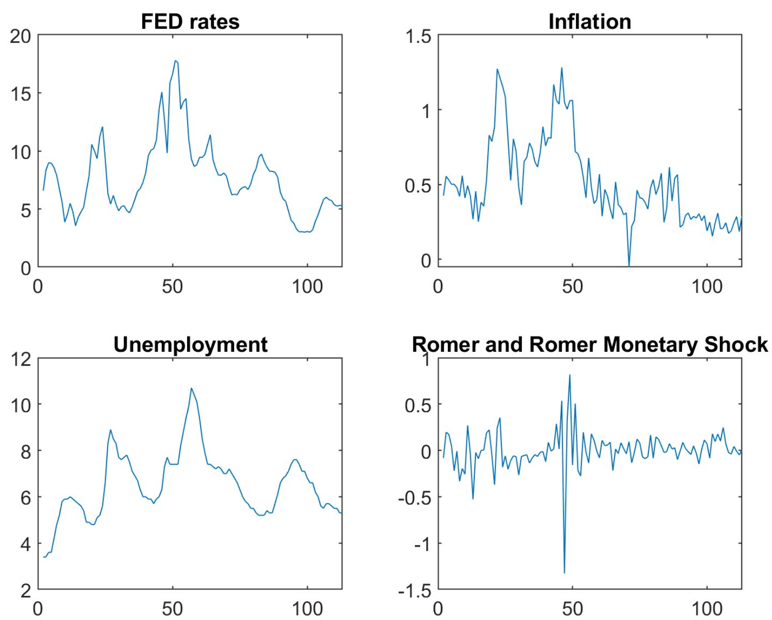

Figure 9: Time series of the variables used in Romer and Romer (2004)

by a simple graphic check we realize that they are exactly the same of the ones presented in the paper; in particular we can notice the spikes in the new variables introduced by Romer and Romer, that corresponds to the years in which the FED decided that the current level of inflation was too high and that it was willing to endure output losses to reduce it.

After this brief introduction we can start with the real analysis. We divide it into two separate sections: the first in which the Granger causality (GC) test is computed by hand with an F-test, and the second in which we use the command "gctest" of the Econometric Toolbox, which allows to compute immediately the Granger causality test. In both cases results are very similar.

F-test for the Granger causality

After having imported the data into the workspace, we construct the data set for the independent variables, which includes all the lags of the four variables, and a separate data set for the dependent variables. Notice that we partition the independent variables in two subgroups (X_1 and X_2), with the first one including all the lags for Inflation ($\mathcal{I}$), Unemployment ($\mathcal{U}$) and the Federal Fund Rates ($\mathcal{F}$) and the second one only including the lags for the variable of interest, the Romer Romer predictor ($\mathcal{R}$). This division will be useful to derive the variance-covariance matrix of the Romer Romer estimators. Notice also that in the matrix of dependent variables (Y), Romer Romer is not included as we are not interested in determining whether its Granger caused by other variables. In particular, our only interest is the joint statistical validity of the four lagged Romer Romer variables when used to predict inflation, Unemployment and the Federal Fund Rate.

We then run the three regressions separetly, exploting a for loop that estimates the OLS coefficients for all the regressors and derives the joint statistical validity of the Romer Romer lags through an F-test. The procedure is as follows. The linear regression below is estimated, where y_t is respectively $\mathcal{I}_t$, $\mathcal{U}_t$ and $\mathcal{F}_t$ depending on the iteration of the cycle.

$$\begin{aligned} y_t = \beta_0 &+ \beta_1\mathcal{I}_{t-1} + \beta_2\mathcal{U}_{t-1} + \beta_3\mathcal{F}_{t-1} \\ &+ \dots \\ &+ \dots \\ &+ \dots \\ &+ \beta_{13}\mathcal{R}_{t-1} + \beta_{14}\mathcal{R}_{t-2} + \beta_{15}\mathcal{R}_{t-3} + \beta_{16}\mathcal{R}_{t-4} \\ &+ \epsilon_t \end{aligned}$$

Notice that the dots replace the second, third and fourth lag of the three variables.

We then compute the residuals of the the regression and use them to derive the Sample variance of the error term (s^2). Then we proceed by deriving the variance covariance matrix of Romer Romer estimates. Here, the previous partition of the independent variables becomes useful since as the sample var-cov matrix of the Romer Romer estimates can be written as:

$$\widehat{Var(\mathbf{b_2})} = s^2(X_2' m(X1) X_2)^{-1}$$

Where $\mathbf{b_2}$ is a vector composed of the four estimators: β_{13}, β_{14}, β_{15} and β_{16}. Notice that the Orthogonal projector $m(X_1)$ is by definition equal to:

$$m(X_1) = I_n - P(X_1) = I_n - X_1(X_1'X_1)^{-1}X_1'$$

Which can be easily computed. After having derived $m(X_1)$ and the resulting sample var-cov matrix for the $\mathbf{b_2}$ estimates, we can easily compute the F-statistic given the Null hypothesis: $H_0: \quad \beta_{\mathbf{2}} = 0:$

$$F - stat = \frac{b_{\mathbf{2}}'(\widehat{Var(b_{\mathbf{2}})})b_{\mathbf{2}}}{k}$$

Where k is the number of restrictions in test, which in this case is equal to four (notice that imposing $\beta_2 = 0$ is equivalent to imposing $\beta_{13} = \beta_{14} = \beta_{15} = \beta_{16} = 0$ which imposes four restrictions). The table below summarizes the results:

Summary			
Romer Romer lags	Inflation	Unemployment	Federal Fund Rate
RR_{t-1}	-0.1580	0.2424	1.5184
RR_{t-2}	-0.2698	0.3979	-0.9740
RR_{t-3}	-0.1946	0.1490	0.5142
RR_{t-4}	-0.0075	0.3231	-1.7606
F-statistics	2.1295	2.1878	9.0610

Table 1: The table shows the OLS estimators for the four lags of the Romer Romer variable, and the their joint statistical significance through the resulting F-statistics.

From Table 1 we observe that the Romer Romer predictor is mildly statistically significant for both Inflation and Unemployment, with the F-statistics being greater than the critical value at the 90% level (2.0077) and below the 95% critical value (2.4718). In essence, lagged values of Romer Romer are not a strong predictors for the current level of Inflation and Unemployment. On the other hand, there is enough evidence to suggest that the current level of the Federal Fund rate is Granger caused by the Romer Romer regressors, with an F-statistics that is well above the 99% critical value (3.5325). These results are in line with findings in Romer & Romer (2004).

"gctest" from Econometrics Toolbox

In this section we present the more immediate and time-saving method to compute the GC test. In particular we take advantage of the command "gctets" provided by the Econometric Toolbox that we installed. But before running the test it is necessary to build up the correct VAR process with proper functions. Firstly we read the Excel file and identify each column as a variable with the command "readvars" (the names of the 4 variables are now "date", even if useless, "infl", "unemp", "ffr", "Romer". See the matlab file for further details). Then, we use the function "varm" to creat a VAR proces (called "VAR" in the file) with 4 series and 4 lags as input; indeed we need the 4 variables identified above and it is required to use 4 lags.
Notice that "VAR" now is an object with many variables inside, but for the moment it is empty; indeed this model needs to be calibrated with the data from the Excel file.
In order to do this we use the command "estimate" using as inputs "VAR" and the 4 variables of interest, i.e. "infl", "unemp", "ffr" and "Romer" (we also collect the names of the series with the variable "varnames".
We then summarize our estimated var, called "EstVAR" (see the file).
Now we are ready for the GC test: the Econometric Toolbox has a command called "gctest" that allows to obtain the results for a GC test; in particular the default option id a χ^2 distribution, and a test for single causality of each variable. Moreover it has the "leave-one-out" default option, that means that the test takes one variable at the time and test for the GC conditioned on all the other variables in the model. We call "h = gctest(EstVAR)" in our file and we obtain the following result:

H0	Decision	Distribution	Statistic
"Exclude lagged UNEMPLOYMENT in INFLATION equation"	"Cannot reject H0"	"Chi2(4)"	4.33
"Exclude lagged FFR in INFLATION equation"	"Cannot reject H0"	"Chi2(4)"	5.9699
"Exclude lagged ROMER in INFLATION equation"	"Cannot reject H0"	"Chi2(4)"	8.5178
"Exclude lagged INFLATION in UNEMPLOYMENT equation"	"Cannot reject H0"	"Chi2(4)"	7.755
"Exclude lagged FFR in UNEMPLOYMENT equation"	"Reject H0"	"Chi2(4)"	9.6494
"Exclude lagged ROMER in UNEMPLOYMENT equation"	"Cannot reject H0"	"Chi2(4)"	8.7512
"Exclude lagged INFLATION in FFR equation"	"Reject H0"	"Chi2(4)"	12.212
"Exclude lagged UNEMPLOYMENT in FFR equation"	"Reject H0"	"Chi2(4)"	19.553
"Exclude lagged ROMER in FFR equation"	"Reject H0"	"Chi2(4)"	36.244
"Exclude lagged INFLATION in ROMER equation"	"Cannot reject H0"	"Chi2(4)"	2.3428
"Exclude lagged UNEMPLOYMENT in ROMER equation"	"Cannot reject H0"	"Chi2(4)"	3.4774
"Exclude lagged FFR in ROMER equation"	"Cannot reject H0"	"Chi2(4)"	5.2552

Figure 10: GC test with a χ^2 distribution and "leave-one-out" option

as we can see, here there are many outputs, but we are interested only in the ones that explain if there is GC causality of the Romer variable, i.e we're interested in rows 3,6 and 9 of this table. As we can see we reject in only in the case of the effect of "ROMER" on "FFR"; in reality if we take a closer look at the p-values of this test we realize that the rejection also for inflation and unemployment is really borderline (the rows are the same of before):

PValue	CriticalValue
0.36318	9.4877
0.20141	9.4877
0.074349	9.4877
0.10098	9.4877
0.046767	9.4877
0.067628	9.4877
0.015841	9.4877
0.00061181	9.4877
2.5777e-07	9.4877
0.673	9.4877
0.48132	9.4877
0.2621	9.4877

Figure 11: p-vales for the χ^2 test

it's clear that the effect of "ROMER" on "FFR", in Grenger-causality sense, is very strong, and indeed the corresponding p-value is very small.
But at the same time, if we give a closer look at the p-values for inflation and unemployment, we notice that they're not so far from the critical value of 5%, so the decision of rejecting or not the null is really borderline. This situation is consistent with the one we studied before with the "manual"approach.

Moreover we can give a more correct and complete view by running the "gctest" specifying the 'Test' "f" option. The procedure is always the same: the command (now the result saved in the variable "h1") runs a GC test but considering an F distribution instead of the χ^2 used as a default option. The resulting table is the following:

H0	Decision	Distribution	Statistic
"Exclude lagged UNEMPLOYMENT in INFLATION equation"	"Cannot reject H0"	"F(4,91)"	1.0825
"Exclude lagged FFR in INFLATION equation"	"Cannot reject H0"	"F(4,91)"	1.4925
"Exclude lagged ROMER in INFLATION equation"	"Cannot reject H0"	"F(4,91)"	2.1295
"Exclude lagged INFLATION in UNEMPLOYMENT equation"	"Cannot reject H0"	"F(4,91)"	1.9387
"Exclude lagged FFR in UNEMPLOYMENT equation"	"Cannot reject H0"	"F(4,91)"	2.4123
"Exclude lagged ROMER in UNEMPLOYMENT equation"	"Cannot reject H0"	"F(4,91)"	2.1878
"Exclude lagged INFLATION in FFR equation"	"Reject H0"	"F(4,91)"	3.053
"Exclude lagged UNEMPLOYMENT in FFR equation"	"Reject H0"	"F(4,91)"	4.8882
"Exclude lagged ROMER in FFR equation"	"Reject H0"	"F(4,91)"	9.061
"Exclude lagged INFLATION in ROMER equation"	"Cannot reject H0"	"F(4,91)"	0.58569
"Exclude lagged UNEMPLOYMENT in ROMER equation"	"Cannot reject H0"	"F(4,91)"	0.86936
"Exclude lagged FFR in ROMER equation"	"Cannot reject H0"	"F(4,91)"	1.3138

Figure 12: GC test with an F distribution and "leave-one-out" option

as we can see, also in this case we reject the null for "ROMER" on "FFR", but not for the other variables (always look at rows 3,6 and 9 of this table). Giving a closer look to the p-values of the tests:

PValue	CriticalValue
0.36991	2.4718
0.21102	2.4718
0.083482	2.4718
0.1107	2.4718
0.054679	2.4718
0.076535	2.4718
0.020735	2.4718
0.0012896	2.4718
3.3133e-06	2.4718
0.6738	2.4718
0.48555	2.4718
0.27082	2.4718

Figure 13: p-vales for the F test

we easily recognize the low level of the p-value for the "ROMER" variable, while the value for "INFLATION" and "UNEMPLOYEMNT" is borderline. Lastlty, the F-statistics we obtain with this test are the same we have obtain manually, which is, we think, a proof of the correctness of our results.

Just for the sake of completeness we mentioned in our file the test ("k" and "k1" respectively for the χ^2 and the F tests) that tests the null hypothesis that all other response variables do not jointly Granger-cause response variable. This tests is a joint test for Granger causality, and as such it is not really useful for our purposes.

Point 4

The first thing we do for this point is to study the DGP, i.e.:

$$\begin{bmatrix} x_t \\ y_y \end{bmatrix} = \begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \eta_t \end{bmatrix}$$

We are asked to:

- Generate 500 observations from the given VMA form.
- Estimate a VAR with 4 lags on the data generated
- Retrieve the structural shocks from the estimated VAR.
- Retrieve the VMA form and compute the empirical Impulse response functions.
- Repeat the first four steps in a Monte Carlo fashion.
- Compare the Empirical and Theoretical IRFs.

In the Matlab file we begin by setting the initial parameters that use in the DGP and creating the empty vectors that will store the four impulse response functions. The four beginning items in the list above will be then conducted inside the for loop (indexed by s).
We begin with the DGP. The structural shocks (η_t and ϵ_t) of the process are randomly generated according to the given var-cov matrix and saved in u. each realization is then computed according to the expression below:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{\beta}{1-\beta} \end{bmatrix} \eta_t + \begin{bmatrix} 0 \\ \frac{\beta^2}{1-\beta} \end{bmatrix} \epsilon_t + \begin{bmatrix} 0 \\ \beta \end{bmatrix} \epsilon_{t-1} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \epsilon_{t-2}$$

After having generated 500 observations through the for loop we move to the VAR estimation. The matrix of regressors is composed using another loop which stores at each cycle a new lag of x_t and y_t in the matrix X. The loops ends when four lags are present in X. Notice that, in order to have X and Y of the same length, we reduce the sample size to 496. We then run the regression, for which we save the OLS estimates and the variance of the residuals respectively in the variables **b** and **var res**. We then move to the derivation of the structural shocks using a Choleski identification scheme. It can be proved that any variance covariance matrix (by property of being positive definite) can be decomposed as the product of a lower triangular matrix G and its transpose. The function *chol( )* applied to the estimated variance covariance matrix of the VAR residuals returns the lower triangular matrix G. The importance of the matrix G (or Choleski factor) relates to its connection with the underlying structural form of the VAR. In particular it can be shown that:

$$
\begin{aligned}
Y_t &= C(L)Y_{t-1} + v_t \\
G^{-1}Y_t &= G^{-1}C(L)Y_{t-1} + G^{-1}v_t \\
AY_t &= B(L)Y_{t-1} + u_t
\end{aligned}
$$

So, through the Choleski factor we can derive the SVAR (last equation) from the estimated VAR (first equation) which is just its reduced form. From the equations above we clearly see that the relation between the residuals and the structural shocks is: $u_t = G^{-1}v_t$. In the file the structural shocks are saved in the variable "shocks".

An important remark to make is that our estimated VAR, or reduced form of the SVAR, it's not linked one to one to the structural form computed through the Choleski factor. Actually, for any reduced form we could potentially derive infinitely different consistent structural forms (and therefore many different structural shocks) depeding on the restrictions we impose on the SVAR. The Choleski identification scheme identifies the SVAR with a lower triangular matrix A and with shocks having an identity var-cov.

The next step is to derive the Impulse response functions (IRFs). Before going through the empirical derivation, let us identify the theoretical IRFs.

Going back to the DGP, we notice that the process can be seen as a VMA since we have x_t, y_t expressed as a function of the error terms ϵ_t, η_t. The useful thing of this representation is that we can see how the DGP behaves when we give unitary shocks to the error terms; if we suppose to observe a unitary shock to ϵ_t, η_t at time t we can see how the process evolves in the subsequent periods of time.

So now we're going to present what happens to the matrix of this VMA process at each t after this unitary shock.

t:

In this case we have that only the contemporaneous error terms will be affected, so all of the lags components in the matrix can be neglected; it follows that the first period shock can be represented in matrix form as follows:

$$
\begin{bmatrix} 1 & 0 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} \end{bmatrix}
$$

$t+1$:

When we study the transmission of the shock in a VMA process we assume to give an impulse only at t, and then the leads of the error terms will be no more affected, for this reason the only component of the matrix that will store some effects in this situation is the one with only one lag, i.e. βL (that indeed multiplies in the extended form ϵ_{t-1}). It follows that the second impulse will be:

$$\begin{bmatrix} 0 & 0 \\ 0 & \beta \end{bmatrix}$$

$t+2$:

The reasoning we follow is exactly the same, and it's easy to see that in this case the only component left with some effect from the original shock is the one with two lags, i.e. L^2. So the matrix will be:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

From this moment onward the effect of the original impulse vanishes, and so all the other components of the theoretical IRF will be:

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Notice each Impulse response (2x2 matrices) has the same structure. The upward left element is the Impulse response (IR) of x_t to a shock in η while the upward right is the IR to a shock in ϵ. The same logic applies to the two downward elements but the IRs are for y_t. For this reasons we are able to derive four different IRFs, each of them describing the IR of a different variable to a different shock.
As already seen the VMA form is useful for the derivation of the IRFs. To derive the VMA form of our VAR we need to exploit its Companion form which is the VAR(1) representation of the model. From the VAR(1) representation the MVA becomes much easier to derive. First, The companion form is derived as follows:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} x_{t-2} \\ y_{t-2} \end{bmatrix} + \dots + \boldsymbol{v_t}$$

$$\longrightarrow \boldsymbol{z_t} = A\boldsymbol{z_{t-1}} + B\boldsymbol{z_{t-2}} + C\boldsymbol{z_{t-3}} + D\boldsymbol{z_{t-4}} + \boldsymbol{v_t}$$

$$\longrightarrow \begin{bmatrix} \boldsymbol{z_t} \\ \boldsymbol{z_{t-1}} \\ \boldsymbol{z_{t-2}} \\ \boldsymbol{z_{t-3}} \end{bmatrix} = \begin{bmatrix} A & B & C & D \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{z_{t-1}} \\ \boldsymbol{z_{t-2}} \\ \boldsymbol{z_{t-3}} \\ \boldsymbol{z_{t-4}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{v_t} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \boldsymbol{0} \end{bmatrix}$$

$$\longrightarrow \boldsymbol{q_t} = M\boldsymbol{q_{t-1}} + \boldsymbol{w_t}$$

Notice that the elements in A, B, C and D are all the OLS estimates that have been computed through the VAR. Since the Companion form can be seen as a VAR of order one (as in the last equation), its VMA form is known at takes the following form:

$$\boldsymbol{q_t} = \boldsymbol{w_t} + M\boldsymbol{w_{t-1}} + M^2\boldsymbol{w_{t-2}} + M^3\boldsymbol{w_{t-3}} + \dots$$

From this form we are interested in retrieving the VMA form of the vector z_t, which is the first element of the vector q_t. We can easily do this by considering just the first row of the RHS in the equation above. Notice that the only non zero value in w_t is in the first position and its the vector v_t, meaning that when computing the first row we'll get an infinite summation of the top left matrix in M^n, which is a 2x2 matrix, and v_{t-n} (e.g for n equal to 1 we get the term Av_{t-1}). Let us denote with C_n the 2x2 matrix in the top left cells of M^n, so that we rewrite the VMA of z_t as:

$$\boldsymbol{z_t} = \boldsymbol{v_t} + C_1\boldsymbol{v_{t-1}} + C_2\boldsymbol{v_{t-2}} + \dots$$

By recalling that $G\boldsymbol{u_t} = \boldsymbol{v_t}$, we can substitute v_t to get a function of the structural shocks:

$$\boldsymbol{z_t} = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = G\boldsymbol{u_t} + C_1G\boldsymbol{u_{t-1}} + C_2G\boldsymbol{u_{t-2}} + \dots$$

Where the sequence of C_nG are the Impulse response of z_t to the structural shocks in u. Going back to the Matlab file, after having defined the Companion matrix (C) we compose 4 vectors (IRF_i)that will store the four IRs. A for loop is then generated where at each cycle a new C_n is computed and multiplied with G to get the IRs matrix, the elements of the matrix are then divided into each of the four IRF_i (notice we let the cycle have 496 repetitions). The final step is to save the four vectors in the $IRF_i storing$ matrix. At the end of the Monte Carlo repetitions of the steps we have just described we should have four matrices containing a sample of 300 IRFs.
After having extracted the means and the confidence intervals for each IRF we can then compare it with the Theoretical one. The plots below illustrates our results.

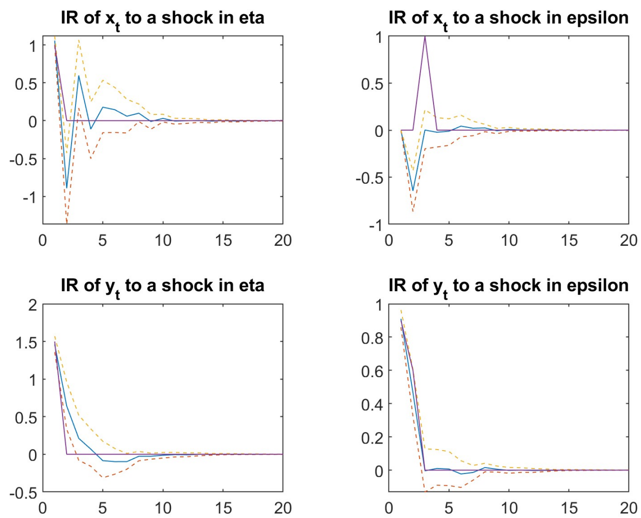

Figure 14: IRFs: Theoretical vs Empirical

As we can see, the comparison between theoretical and empirical IRFs brings quite different results; they are not dramatically different, but there are some discrepancies between the two.

The most evident difference is the one in the top-right of our last picture: here we can see that after 0 we observe a clear negative spike in the empirical IRF, that doesn't match the theoretical one; we tried to explain this by taking into account the lags on both the DGP and the VMA representation of the VAR(4) that we have estimated.
This picture in particular shows the IR of x_t to a shock in ϵ: if we go back to the DGP, we easily see that the only way in which ϵ affects x_t is by its effect at time $t-2$, but it has no other lags; on the other hand, while written with the companion form and the VMA representation, the original VAR is expressed with (potentially) infinite lags so it will store the effects of ϵ and of all its lags. The empirical shock is negative due to the OLS estimate. If we compare the theoretical and empirical IRFs we see that the initial shock is present in both graphs; however the one on the theoretical is equal to one and due to the shock on $t-2$, while the shock on the empirical is lower because of the influence of the other lags, which are not present in the theoretical.
Notice also that at time 0 the restrictions of the Choleski match the theoretical model, due to lower-triangularity.

A similar reasoning can be applied to the top-left picture; here we observe the effects of η on x_t.
The theoretical is affected only by the contemporaneous shock. The empirical

IRF matches the shock at time zero and then adjusts the remaining lags oscillating around zero.

The two bottom pictures are very similar: both of them are almost coincident with the theoretical IRF. We think this is because both ϵ and η affects y_t contemporaneously; moreover the IRF for ϵ is almost perfect since the theoretical has not only the effect at time t, but also the one at time $t-1$. Of course, after the first lags both converge to zero creating even a clearer overlapping picture.

To sum up this explanation, we think that we have some problems of identification in our DGP: it has many zeros on the lags, a property that is not particularly indicated when working with SVAR, as pointed by Sims. An important remark is that the IRFs of y_t performs better than the one for x_t. The variable y_t enjoys a contemporaneous effect from both shocks, meaning that when a shock is observed, the OLS of the VAR captures better their relative weights. On the other hand, in the case of x_t, we don't observe any contemporaneous effect, so the OLS doesn't capture this property, and the empirical IRF takes more time to adjust to the steady state.
One last interesting thing to notice is that, as shown by Ronayne (2011), when the VAR coincides with the DGP, then the procedure is optimal for all time horizons; instead, when the DGP and the estimated VAR do not coincide, the resulting IRF could be biased. We think also this kind of reasoning can be applied to our model, given that we start with a DGP and compute the theoretical IRF, while the empirical one is based on a VAR process.

References

- Bruno, Giovanni (2022), "Econometric Models: Lecture Notes for 2021/2022 ESS"
- Ghysels, Eric Marcellino, Massimiliano (2018), "Applied Economic Forecasting Using Time Series Methods", Oxford University Press
- Enders, Walter (1995), "Applied Econometric Time Series", Wiley Series in Probability and Mathematical Statistics, John Wiley Sons
- Hamilton, James D. (1994), "Time Series Analysis", Princeton University Press
- Romer, Christina D. Romer david H. (2004), The American Economic Review, Vol. 94, No. 4, pp. 1055-1084
- Ronayne, David (2011), "Which Impulse Response Function?", WARWICK ECONOMIC RESEARCH PAPERS
- Sala, Luca (2022), "Notes on Time Series Analysis"
- Sims, C.A., Stock, J.H. and Watson, M.W. (1990) Inference in Linear Time Series Models with Some Unit Roots. Econometrica, 58, 113-144. https://doi.org/10.2307/2938337