Macroeconometrics Problem Set 3: Explanatory Notes

Corti Pietro (3061891), Reichlin Vito (3081010)

27 November 2022

Point 1

The first thing we do is to upload the dataset from the Excel file "data_ps3" using the command "readvars" and identifying the variables "time", "GDP", "PD" (Price Deflator) and "FFR" (Federal Fund Rates).
Given that the quarters ("time") is not a variable of interest we delete it and consider only the last three variables, collected in the new "data" variable.

The second step of our analysis is to prepare the original estimation that we will use in the bootstrap procedure; to do this we identify the dependent and independent variables, namely "Y0" and "X0", containing all the three datasets imported before.
To simplify the notation and the code, we decide to implement a $VAR(1)$, but a similar procedure can be implemented using any other number of lags. Of course our case is simpler given that there is no need for a companion form, but when $p > 2$ in a $VAR(p)$ to obtain the VMA representation and the IRFs, the companion form of the estimated matrix is needed.
Using a simple OLS estimate we then obtain the original matrix of coefficients "b0".

The third section is relatively simple given that we simply collected the (empty) vectors for the future IRFs and the forecast error variance decomposition (the "FVD"s in our file).

In the following section of our Matlab file we start with the loop of 1000 simulations in order to obtain the bootstrap. Implicitly, we follow the steps suggested in the hint of the document by creating a vector of integers "r" that we are going to use for the bootstrap with replacement.
The loop for this command, within the bigger loop for the bootstrap, basically extract one row of the residuals and use to create a new vector; then "bs" will become the new bootstrapped residual that we can use in our new estimation of the dependent variable. We cannot use "permute" because it would switch rows with columns, modifying the order of the original residual vector, and so the replacement will be compromised.

Once we have done this we are going to use the new "bs" vector to compute the value of the new dependent variable "Y" being simply the sum of previous dependent variable "X0" and the new residual "bs". Repeating this process will bring to 1000 different observations.
The following steps are similar to what we have done in previous Problem Sets, i.e. we compute the variance of the residuals "var_v", implement the Choleski "G" and use them to compute its inverse "A", the structural shocks "u" and the IRFs.
These ones are obtained with a loop in which we multiply the powers of the matrices for the OLS (called now "C") with the Choleski "G".
The last part of the loop is dedicated to the computation of the Forecast Error Variance decomposition (FEVD); to explain it let's consider a VMA process like the following one:

$$y = \Psi(L)\epsilon_t$$

the one-step forecast error is given by the following formula:

$$Var(FE_1) = Var(y_{t+1} - y_{t+1|t}) = \Omega$$

and the two-steps ahead forecast is given by:

$$Var(FE_2) = Var(y_{t+2} - y_{t+2|t}) = \Omega + \Psi_1 \Omega \Psi_1'$$

we can proceed in this way and continue to explain the decomposition of the variance. So, given that we are applying a Choleski decomposition,that means $\Omega = GG'$, the general formula for the FEVD will be:

$$Var(y_{t+1} - y_{t+1|t}) = \Omega = GG' = G_{(1)}G_{(1)}' + G_{(2)}G_{(2)}' + ... + G_{(N)}G_{(N)}'$$

this is what we have done in the loop in our Matlab file, using the fact that dealing with a $VAR(1)$ we were able to move to the VMA representation by using the power of the matrix "C" (see the Matlab file for more details). The decomposition that we have obtained is visible in the variables "FV" for the total decomposition, "fv" for the effect of each variable in explaining the monetary shock and their ratios "fvd".

Lastly, after these computations, we ended up with 1000 IRF and we computed the means and the confidence intervals obtained with the bootstrap method ("m_IRF and "ci_IRF1" respectively in our file). The graphical representation is the following:

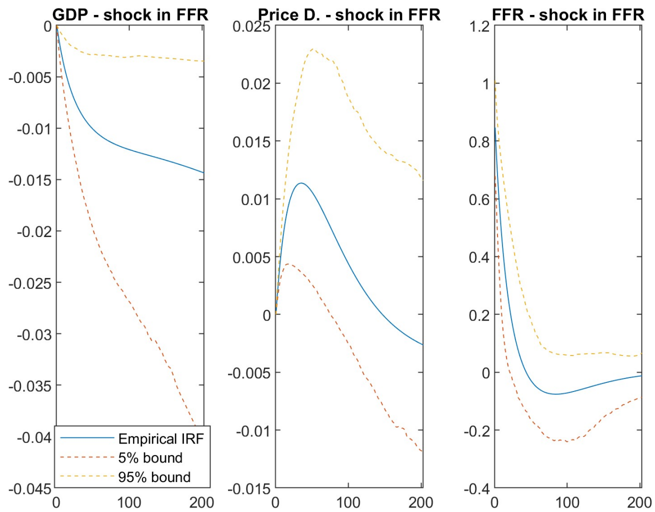

Figure 1: IRF for the monetary shock on the different variables: GDP, Price deflator and FFR

as we can see the picture seems to reflect the theoretical predictions given by economic theory; after the monetary shock GDP decreases but then it seems to stabilize (here it is not clearly visible given that the figure is truncated after 203 observations which is the length on the data we have).
It also seems very probable to observe the so called ”prize puzzle” because after a monetary tightening it’s counter intuitive to observe increasing inflation; this phenomenon is well know in the literature and Sims (1992) proposed that it is possible to correct it by adding commodity prices.
Lastly the FFR decrease before stabilizing after the shock, as shown in the third panel.

Point 2

The first thing we do here is to clear the file and import the new data on the second sheet of the Excel file "data_ps3". We divide them in productivity "YL" and hours "HOURS". However, it is necessary to obtain the log of the difference for these variables, given that this is the choice of Galì in his paper. As a result we end up with "deltaYL" and "deltaH"; they are multiplied by 100 to simplify the graphical representation. For the sake of completness we also reported an optional graph with the row data.

The second step is to estimate the VAR. To better explain this more complex point we decided to run a $VAR(4)$ by considering the possible effect of many lags on our variables of interest. To simplify the code we did this by using a "varm" command from "Econometric Toolbox", specifying two series ("deltaYL" and "deltaH") and 4 lags. Then we estimated the empty "varm" with our time series obtaining the variable "EstVAR".
To compute the bootstrap method we also prepared the variables "X0" and "Y0" containing the two series and, with the coefficients saved from "EstVAR" above, we created the matrix of OLS "B0"; this shift from one method to the other was needed to compute both the IRF and the bootstrap. The last thing we did before the implementation of the loop for the bootstrap is the creation of the variables of residuals "v0".

In this section of our code we implemented the loop for the bootstrap, in a similar fashion of the previous point; iterations are equal to 1000 and the method to estimate the "extracted - with replacement" row residual is the same of before (always "randi" for the vector of integers, see Point 1).
Moreover also the generations of new variables is very similar to what was done before; in this case we prepare the DGP (Data Generating Process) for the new $VAR(4)$ by using the variables "Y", "X" and bootstrapped residual "bs", with these data a new vector auto regression is implemented with a new "varm". The subsequent loops were needed because, using objects like the matrix "B(2,2,2)" that are not simply matrices but multi-dimensional objects, the position of the coefficients on the main diagonal of OLS matrix is inverted. Concluded these adjustments we obtained the object "B" which contains all the OLS of the two variables for each different lag.

The loop continues by setting down the variance-covariance matrix "V" and the companion form for each matrix in "B", which is called "C". This matrix is needed in order to get the VMA representation for the computation of the IRFs. In fact let's remember that given, a VAR of this kind:

$$y_t = AY_{t-1} + \epsilon_t$$

we can move to the VMA representation, i.e.

$$y_t = C(L)\epsilon_t$$

by considering the following relationship between matrices:

$$C_i = A^i$$

These matrices are then collected in the object "W" which has dimensions 2X2X500 (we approximated 500 iterations for the VMA).

It is worth studying in depth now the restriction we made to estimate our model, which are the ones considered also by Galì (1999). Considering for simplicity a $SVAR(1)$ process, i.e.

$$Az_t = Bz_{t-1} + u_t$$

and its reduced form

$$z_t = CZ_{t-1} + v_t = A^{-1}Bz_{t-1} + A^{-1}u_t$$

the VMA representation is

$$z_t = D(L)v_t = (I + CL + C^2L^2 + ...)v_t$$

if it is evaluated at $L = 1$ we obtain the result of a geometric series:

$$D(1) = (I - C)^{-1}$$

from this we can obtain the long-run restriction; however to simplify our code we used the process based on the Choleski factor, where we computed the matrix:

$$D(1)\Omega D(1) \quad where \quad \Omega = Cov(v_t)$$

whose Choleski decomposition id given by SS'.
We then computed $K = D(1)^{-1}S$, and obtained the IRFs from:

$$D(L)K$$

This is exactly the approach that can be seen in our Matlab file with more detail; the variable "D1" (corresponding to $D(1)$ in the theoretical description) has been obtained by the sum of the different VMA matrices (up to 500 as said) collected in "W".
So the IRF have been obtained by the loop in which we multiply "W" by the variable "K" as identified in the theoretical description above.
The last step was to separate the IRF concerning respectively "deltaYL" and "deltaH" by using the command "squeeze" (since again, after the loop the IRF is a 2X2X500 object, which is not graphically representable). The IRF for the GDP is simply given by the sum of the previous two, as expresses in the paper.

After the correct identification of the technological shock and the non-technological shock, we represented the IRF without bands. The result id the following:

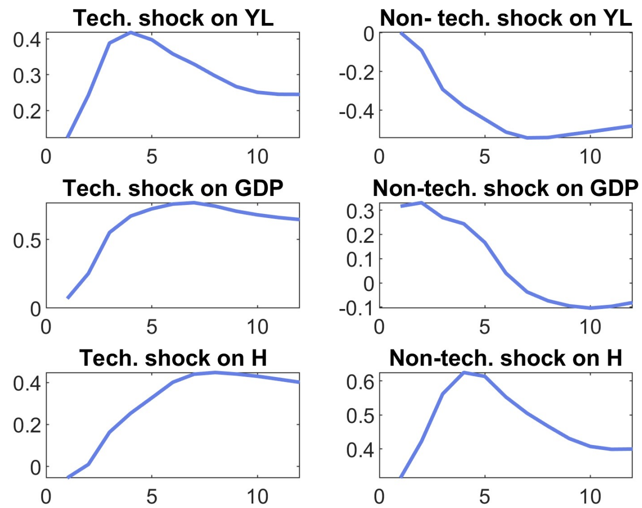

Figure 2: Estimated IRFs from the bivariate $VAR(4)$

as we can see the graphs are very similar to Galì's one, even if there are some slightly difference; this could be given by the fact that the data, as said in the text, are not exactly the same. Moreover, reading the original paper it is not very easy to understand how many lags were used in the original model.
Despite these differences we can see that productivity increases initially and then stabilizes after a technological shock; on the other hand, and in line with the model restrictions, the shock vanishes if it is caused by non technological factors.
GDP is pretty much stable after an initial shock if the driving force is technology and it decreases after a shock in case the driving force is not technology.
Lastly, the effect of technology on hours is basically null in the long run, and this is the most important aspect of the study conducted by Galì given that it showed the incorrect assumptions of the RBC model. Eventually, a non technological shock causes a pronounced increase in the hours worked, but than a rapid and constant decrease in the long run.

To conclude, the last part of the code is dedicated to the creation of the median values ("m_IRF" for each shock and each variable) for the IRFs and the creation of the bootstrapped confidence intervals ("ci_IRF" for each shock and each variable). See the Matlab code for further details on this point.
Our final result is the followwing:

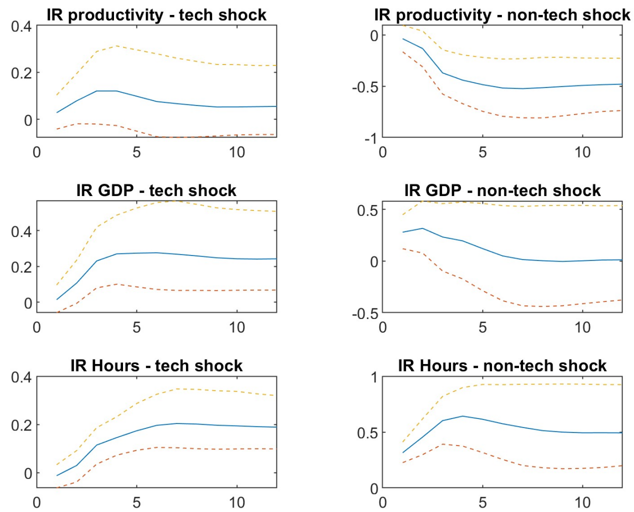

Figure 3: Estimated IRFs from a bivariate $VAR(4)$ with bootsraps

again, the figures are very similar to the one that we have discussed before, but in this case we have considered the median values with their bootstrapped confidence intervals.

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
- Galì, Jordi (1999) The American Economic Review, Vol. 89, No. 1, pp. 249-2
- Hamilton, James D. (1994), "Time Series Analysis", Princeton University Press
- Sala, Luca (2022), "Notes on Time Series Analysis"
- Sims, C.A. (1992), Christopher A. Sims, "Interpreting the macroeconomic time series facts: The effects of monetary policy", European Economic Review,Volume 36, Issue 5, pp. 975-1000, ISSN 0014-2921, https://doi.org/10.1016/0014-2921(92)90041-T.(https://www.sciencedirect.com/science/article/pii/001429219290041T)