Macroeconometrics
Problem Set 2

Nunzio Fallico 3072744
Chiara Locatelli 3070485
Matteo Ticli 3077833

November 2022
Università Bocconi

Exercise 1

a)b)

The data generating process is $y_t = \alpha + y_t + \varepsilon_t$, we run the regression $y_t = \alpha + \rho y_t + \varepsilon_t$. We want to test $H_0 : \rho = 1$ vs $H_1 : \rho \neq 1$ by means of a t-test. We find that the empirical probability of rejecting the null using a 95% confidence interval is 0.0506. The distribution of the t-test can be seen in the following graph:

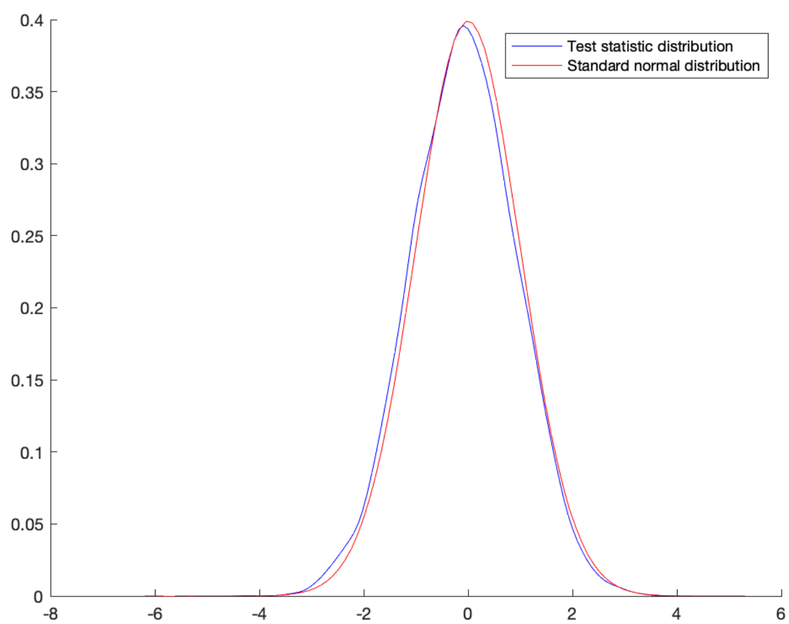

Figure 1: Comparison between test statistics distribution and standard normal distribution

As we can observe, the distribution of the test statistics is very close to the distribution of a standard normal one. As previously seen in Problem Set 1, we would have used the Dickey-Fuller distribution

for the test statistics if the presence of a deterministic regression (α in this case) is neglected i.e. not incorporated in the regression equation. The idea underneath is derived from the paper of Sims, Stocks, and Watson (1990). They have demonstrated, that adding a deterministic trend to the estimated equation, allows the test statistics (t-test of F-test) to be evaluated using the standard normal distribution. The rate of convergence of the parameters is dominated by the parameter with the slowest rate of convergence (the deterministic regressor α in this case).
Moreover, we have displayed the mean of the test statistics which is very close to 0 (mean of the standard normal) it is -0.0783. Again, this is in line with the paper presented by Sims, Stocks, and Watson (1990). We notice that the mean of the distribution of the test statistics, is very close the normal distribution one (0). The interesting part is related to the fact that we can use a t-statistics as statistic for the hypothesis testing without leveraging the one proposed by Dickey and Fuller.

c)

The data generating process now is $y_t = \alpha + \delta t + y_t + \varepsilon_t$, we run the regression accounting for the deterministic time trend $y_t = \alpha + \delta t + \rho y_t + \varepsilon_t$. We perform the same test statistic as in the previous point and we obtain that the empirical probability of rejecting the null hypothesis, using a 95% confidence interval is 0.0486. The distribution of the test statistics can be observed in the following plot:

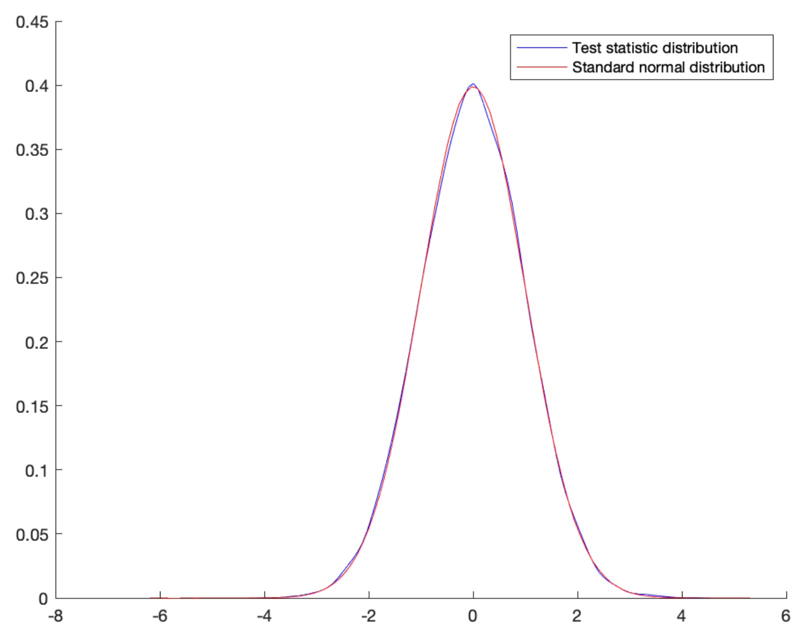

Figure 2: Comparison between test statistics (with time trend) distribution and standard normal distribution

As it can be seen, the distribution of the test-statistics is very close to the standard normal distribution (even more precisely in this case than before). Furthermore, we have displayed the mean of the test statistics that is -0.0021.
The plot and the mean of the empirical distribution lead us to conclude that also in this case the result from Sims, Stocks, and Watson (1990) is still valid as we would have expected.

Exercise 2

Suppose that:

$$\begin{cases} y_t = \phi_y y_{t-1} + \varepsilon_{y,t} \\ z_t = \phi_z z_{t-1} + \varepsilon_{z,t} \end{cases}$$

and we estimate the following regression:

$$y_t = \alpha + \beta z_t + \nu_t \tag{1}$$

Since the two series are completely unrelated, the two data generating processes are substantially different. This reasoning implies that we should expect β equal to 0. We now see what happens if we test $H_0 : \beta = 0$ against the alternative $H_1 : \beta \neq 0$ in different cases, depending on whether either one or both variables are stationary.

Case 1:
Both y_t and z_t are stationary, i.e. $|\phi_y| < 1$ and $|\phi_z| < 1$. By running the regression in (1) and testing the null $H_0 : \beta = 0$, it turns out that with the standard t-test, constructing a 95% confidence interval, we empirically reject the null hypothesis (correctly given the premises) with a probability of 0.0580. Moreover, we observe an R^2 of 0.0016; this value is tends to zero as expected since the two series are unrelated, using one to explain the variation of the other is a useless task and the model will fit poorly. Therefore, when both y_t and z_t are stationary the classical regression model is appropriate.

Case 2:
Suppose that y_t is non-stationary, let us assume it is a random walk, i.e. $\phi_y = 1$, while z_t is stationary, i.e $|\phi_z| < 1$. By running the regression in (1), we now empirically reject the null $H_0 : \beta = 0$ with a probability of 0.1123, while the model's R^2 is approximately 0, more precisely 0.0015. The fit of the model is again very poor given the fact that the two generating process are completely unrelated and estimating one using the other is empirically not appropriate. Moreover, the probability of rejecting the null is now clearly higher than with respect to Case 1 where both processes where stationary, since the standard regression model is not appropriate in this case.

This is because the residuals from regression (1) are non stationary. Assume, without loss of generality that $\alpha = 0$, then:

$$\nu_t = y_t - \beta z_t = \sum_{j=1}^{t} \varepsilon_{y,j} - \beta \sum_{j=0}^{t} \phi^j \varepsilon_{z,t-j}$$

Clearly, the variance of the residuals ν_t will be time-dependent because of the non-stationarity of y_t.

$$Var(\sum_{j=1}^{t} \varepsilon_{y,j}) = t\sigma^2_{\varepsilon_y}$$

The residuals of the regression are thus characterized by a convergent part, however, at the same time the presence of the stochastic component $\sum_{j=1}^{t} \varepsilon_{y,j}$ makes the residuals ν diverge, explaining why the empirical rejection probability is very high even if we are constructing a theoretical 95% confidence interval.

Case 3:
Suppose that both processes are non-stationary but they are integrated of the same order, let us say without loss of generality they are both random walks, i.e. $\phi_y = \phi_z = 1$. By running the regression in (1) and constructing a 95% confidence interval, we empirically reject the null $H_0 : \beta = 0$ with a probability of 0.9266, while the model's R^2 is 0.2404.

This is because the residuals from regression (1) are non stationaty. As usual, assume without loss of generality that $\alpha = 0$, then:

$$\nu_t = y_t - \beta z_t = \sum_{j=1}^{t} \varepsilon_{y,j} - \beta \sum_{j=1}^{t} \varepsilon_{z,j}$$

The variance of the residuals increases as t increases, and the residuals themselves diverge because of $\sum_{j=1}^{t} \varepsilon_{y,j}$ and $\sum_{j=1}^{t} \varepsilon_{z,j}$.
In this case the classical regression model is not appropriate and in order to deal with this kind of scenario, we have to conduct a first difference to rule out non stationary and being able to run the standard OLS regression.

Case 4:
The sequences y_t and z_t are integrated of the same order and the residual sequence is stationary. In this case y_t and z_t are cointegrated. Suppose that both y_t and z_t are the random walk plus noise processes:

$$y_t = \mu_t + \varepsilon_{y,t}$$

$$z_t = \mu_t + \varepsilon_{z,t}$$

where $\varepsilon_{y,t}$ and $\varepsilon_{z,t}$ are white noise, i.e. $, \varepsilon_{y,t} \sim N(0, \sigma^2_{\varepsilon_y})$ and $\epsilon_{z,t} \sim N(0, \sigma^2_{\varepsilon_z})$, and μ_t is the random walk process.

$$\mu_t = \mu_{t-1} + \varepsilon_{\mu,t}$$

By running the regression in (1), we now reject the null $H_0: \ \beta = 0$ with a probability of 1 and the R^2 is 0.9741.

This is because the residuals from regression (1) are non stationary. Assume, without loss of generality that $\alpha = 0$, then:

$$\nu_t = y_t - \beta z_t = \left[\sum_{j=1}^{t} \varepsilon_{\mu,j} + \varepsilon_{y,t} \right] - \beta \left[\sum_{j=1}^{t} \varepsilon_{\mu,j} + \varepsilon_{z,t} \right]$$

The residuals of the regression are then clearly characterized by the presence of a stochastic component $\sum_{j=1}^{t} \varepsilon_{\mu,j}$ that makes the residuals ν_t diverge. This explains why the empirical rejection probability is 1 even if we constructed a theoretical 95% confidence interval.

In conclusion, the classical regression model is not appropriate and to correctly deal with it we should nullify the stochastic trend by considering

$$y_t - z_t = \mu_t + \varepsilon_{y,t} - \mu_t - \varepsilon_{z,t} = \varepsilon_{y,t} - \varepsilon_{z,t}$$

which is stationary.

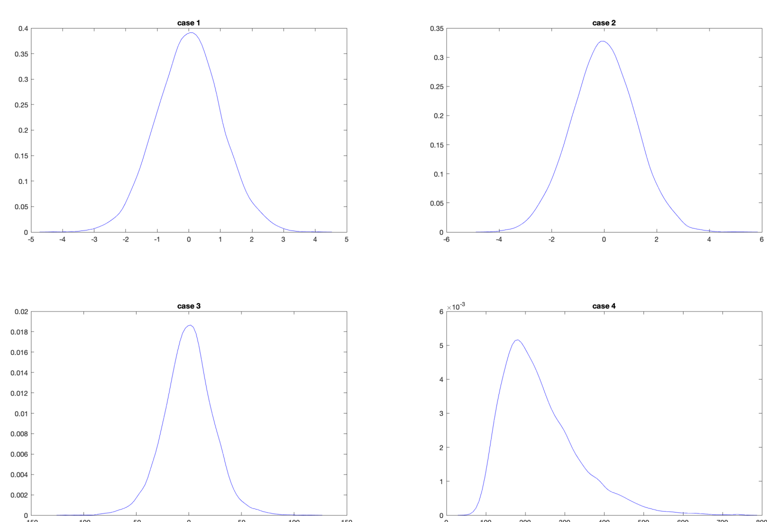

Figure 3: Empirical distribution of test statistic for the four cases of spurious regression considered

Exercise 3

In the data provided are listed four main variables of interest representing US inflation, US unemployment, US federal fund rate and the Romer and Romer index accounting monetary policy shocks. The first three are main macroeconomic variables relative to the US economy, while the fourth is a measure of monetary policy shocks published by Romer and Romer in *The American Economic Review (2004)*, exploited by the authors to overcome the traditional problems of endogeneity and anticipatory movements in the index used to explain other variables of interest e.g. Output and Inflation. The index is able to provide a clear and bias free estimate of the impact of FED's monetary policy decisions: in the case of endogenity, it eliminates much the relationship between interest rates and economic conditions and is able to cover crucial episodes missed by existing indexes (i.e. Bums expansion in early '70s and Volcker stagfaltion in early 80's). Moreover, the Romer&Romer series abstracts from monetary policy actions taken in response by forecasts of economic developing in the months and years to come (i.e purges from the "Greenbook" forecasts provided by FED's analysts relative to the economic development of the near future). Romer and Romer constructed the index in such a way that it can isolate monetary policy shocks, Federal Fund rate changes are only considered if they are the result of the deliberate decisions by the Federal Reserve.

The goal is to test whether the Romer&Romer shock series is able to Granger cause the other three variables of interest (i.e. if R&R is a good forecasting tool for the three other variables mentioned). We estimated a VAR model on the data provided using the econometric toolbox by calling the function *varm(...)*, and then stored the results of the model's estimation. The definition of Granger Causality implies that in the following bivariate VAR

$$\begin{cases} x_t = a_1(L)x_{t-1} + a_2(L)y_{t-1} + \varepsilon_{xt} \\ y_t = b_1(L)x_{t-1} + b_2(L)y_{t-1} + \varepsilon_{yt} \end{cases}$$

the variable x_t does not "Granger cause" y_t if $b_1(L) = 0$. We can therefore test the aforementioned restriction using an F-test, that allows to jointly test the restrictions (we need to test all the coefficients of the specified lags in the model).
Therefore, we want to deploy the Wald Test that is a generalization of the F-test in the multivariate case. In order to perform the Wald Test, we need to be able to compute the Wald Statistics:

$$(R\hat{\beta} - q)'(\sigma^2 R(X'X)^{-1}R')(R\hat{\beta} - q) \sim \chi^2_q$$

The matrix X is composed by the data needed in the format required, and it includes all the lags and specific variables in the appropriate order. The vector R is the auxiliary matrix needed to adjust the restrictions to test with the estimated coefficients from the VAR ($\hat{\beta}$). Finally q is the vector of restrictions that need to be tested and (σ^2) is the Variance Covariance matrix of the error term estimated through the VAR.
Once performed the appropriate hypothesis testing (the null hypothesis is that R&R Granger causes the other three variables), we notice that the the p-value for Inflation and Unemployment is respectively 0.386 and 0.0344 whereas the p-value for Federal Fund Rate is 0.000. Intuitively, it means that we can assess statistical significance that Romer & Romer shocks "Granger causes" the Federal Fund Rate and, depending on which level of significance, Romer & Romer series "Granger causes" inflation and unemployment.

The former situation is clear and the rejection of the null hypothesis cannot be ruled out. The series of monetary shocks are disconnected to interest rates changes in both endogenous responses to economic activities or attempts of forward guidance by policymakers by construction. Therefore, the Romer&Romer series is helpful in forecasting the FFR variable but it does not mean that the Romer&Romer variable is able to cause the FFR. Therefore, we can considered as appropriate, the methodological procedure by which, Romer and Romer derived the estimator that we have just tested. R&R series is created in this peculiar manner, to be used as shock for the FFR thus, forecasting the FFR using the R&R would entail a low forecast error.

The situation is a bit different for the latter two variables. Inflation, as explained in the Romer and Romer (2004) paper, is very hard to capture, especially the identification of the lags required to assess appropriate forecasting. In the paper, there is evidence that some good's prices change after 6 months following a shock captured in Romer&Romer series, while others move after 22 months. There is a an high volatility in measuring this empirical evidence. As far as we are concerned, inflation cannot be entirely "Granger Caused" by the R&R series, using R&R as only measure of forecast is not appropriate. There are some prices that are more sticky than others, some will change more frequently while others will require more time to vary. We conclude that there must be other relevant factors that we are excluding, ex-ante, that would lower the forecast error of Inflation. Moreover, the change in unemployment follows the same exact reasoning. Nonetheless, these two latter variables (Π_t and U_t) are statistically induced by R&R, on the "Granger Causality" base, at 5% significance level. Nonetheless, it would be hard to state that the only factor able to explain the realisation of (Π_t and U_t) is precisely R&R.

Exercise 4

We generate 500 observations from the following data generating process:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & L^2 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} \tag{2}$$

where the $Var(\eta_t) = 1$ and $Var(\varepsilon_t) = 0.8$ while their covariances are 0, implying that the shocks are orthogonal.
Once we have generated the observations for both variables (i.e. x_t and y_t), we need to estimate a VAR model with four lags. We then apply the Cholesky decomposition to the estimated variance-covariance matrix of the VAR model, obtaining the Cholsesky factor G. In order to store the impulse response functions for each simulation, we need to create an all zeros array except for the first entry which is characterized by the presence of the shock (i.e. entry of the considered shock equal to 1) appropriately pre-multiplied by the Cholesky factor G. Considering that $G = A^{-1}$, we are able to move from the reduced form of the model to the structural one. This is inline with the Cholesky identification scheme and it is approproate for the DGP considered.

We appropriately create a nested loop, where we can collect the artificially simulated Impulse Response Function of the first shock η_t on the variables x_t and y_t followed by the second shock ε_t on the same variables x_t and y_t. Since we want to see the impact of the reduced model's shock on the theoretical variables governing the economy, we can use the coefficients estimated from the VAR model with four lags together with the appropriately lagged data, pre-multiplied by the Cholesky factor G in order to isolate the structural form of the model. We end up with a 4D array containing the observations, the variables, the simulations, and the result of the separate shocks. Furthermore, we are able to compute the mean of the simulations, for each observation in time and the 5-th and 95-th percentiles in order to create the confidence band of the observed means. We lastly plot in Figure 4 the theoretical impulse response functions and the empirical ones, together with their confidence bands.

We can finally compare the empirical impulse response functions with the theoretical ones and we find a similar behavior for three out of four of them. The Impulse Response Function of the second shock for the first variable is somehow off with respect to the theoretical one. The theoretical IRF of the second shock on the first variable is null for the first two periods and then is present just for one period, with the shock ceasing to exists at period three already. The empirical IRF, suggests instead, a gradual increase with a peak at period two but a presence of the shock at time one as well. Given our understanding, it seems strange that the first variable is impacted, with two lags, directly by the second shock. We believe that, as nature of the shocks themselves, it is essential that they impact directly their corresponding variable and it follows, at a later point in time, that the shock is reverberated to other variables when the impact of the shocks cannot be disentangled precisely, as the VAR model suggests. Since the second shock impacts variable two at time zero, the VAR estimates, that the same shock would impact variable one at the following period at most and, not being present two periods ahead. We have tried imposing a zero coefficient restriction instead of L^2 and the Impulse Response Functions seems to behave accordingly to the theoretical counterparts as we can observe from Figure 5. Therefore, we consider this to be the main source of error in the DGP specified, to solve the problem, we would impose a zero restriction on the L^2 element of the matrix juxtapposed to the structural shocks with a final form of the DGP that resembles the following:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{\beta}{1-\beta} & \frac{\beta^2}{1-\beta} + \beta L \end{bmatrix} \begin{bmatrix} \eta_t \\ \varepsilon_t \end{bmatrix} \tag{3}$$

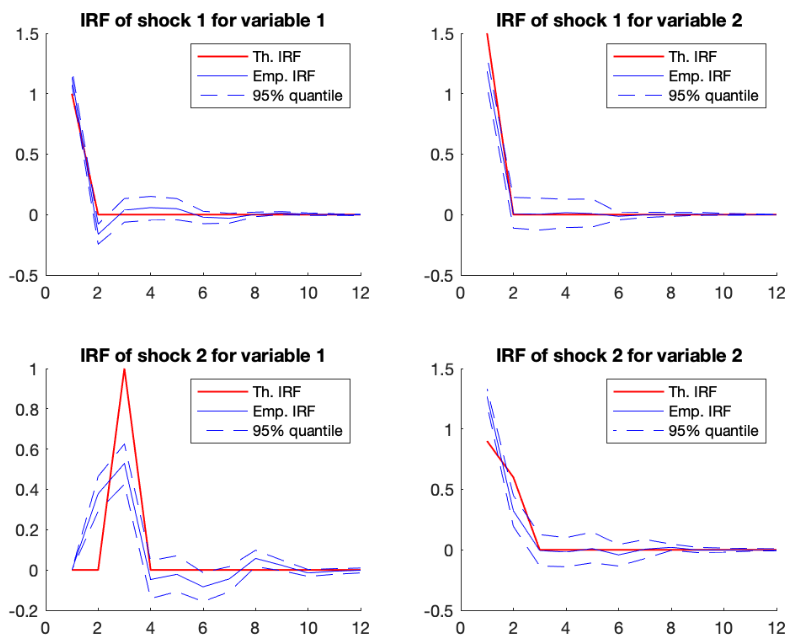

Figure 4: Effects of the exogenous shocks on the model's variables

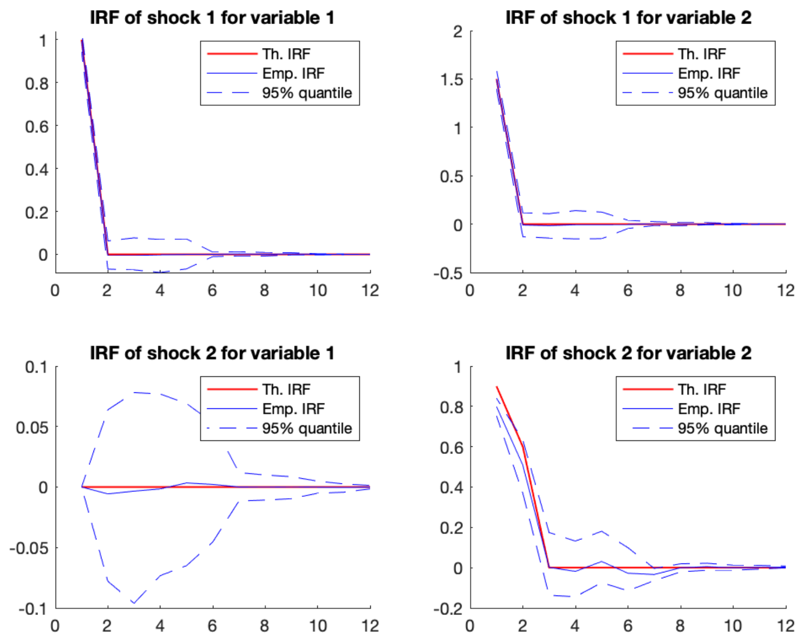

Figure 5: Effects of the exogenous shocks on the model's variables, modified DGP