VAR Models Forecast UK economy

Enric Garcia Paya

Ironhack

enric.garcia.paya@hotmail.com

March 15, 2024

Overview

1. 1 Introduction
2. 2 Data Preparation
3. 3 VAR models, brief introduction.
4. 4 Data transformation
5. 5 VAR model
6. 6 VECM model
7. 7 VARMAX
8. 8 Conclusions

1 Introduction

The 2008 crisis produced an increase of unemployment, debt and adjustments on the public sector throughout cuts on the public spending plus tax rises which favored the impoverishment of the middle classes and with it the rethinking of our society model.

This has been reflected in a radicalization of society and the rise of new extreme left and right parties who openly question the model of society that we have enjoying since the Second World War.

These tensions have put the focus on the ability of economists to predict economic crises or design policies capable of softening their consequences in society.

1 Introduction

The idea is to compare the predictions made by three models and see which of the three best captures the behavior of a set of macroeconomic variables with the data coming from "A millennium of macroeconomic data" from the Bank of England, for the period 2010-2016.

- VAR (Vector Autoregresive model)
- VARMAX (Vector Autoregresive model with a Moving Average Component)
- VECM (Vector Error Correction Model)

2 Data Presentation

As it is said before the data we are going to use is "A millennium of macroeconomic data" from the Bank of England. This dataset contains a broad set of macroeconomic and financial data for the UK stretching back in some cases to the C13th and with one or two benchmark estimates available for 1086.

After selecting our variables and and drooping null values we end up with a data sample with 10 variables annualized or the period 1922-2016 on UK.

2 Data Presentation

- Real GDP 1922-2016
- Consumption 1922-2016
- Investment 1922-2016
- Employment 1922-2016
- Oil Prices 1922-2016
- Public Sector Debt 1922-2016
- TOT 1922-2016
- Total Managed Expenditure 1922-2016
- M1 1922-2016
- House Price 1922-2016

2 Data Presentation

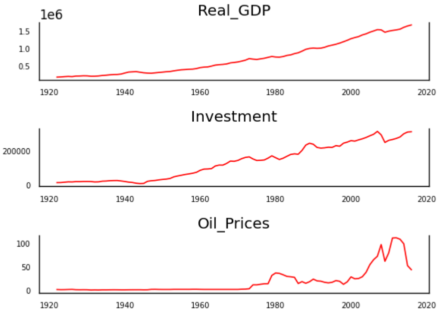


Figure: Figure 1

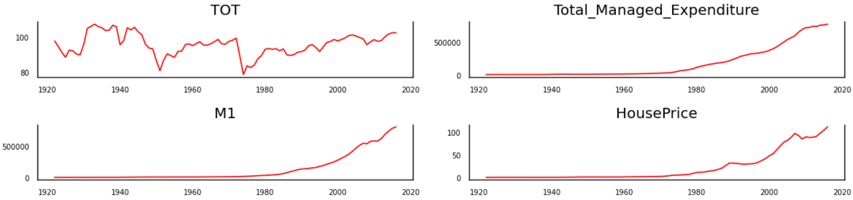


Figure: Figure 2

3 VAR models, brief introduction.

A VAR model is a model made of a vector of stationary variables which his behavior is gonna be explained by its behavior on the past.

$$X_t = A_0 + A_1 X_{t-1} + e_t$$

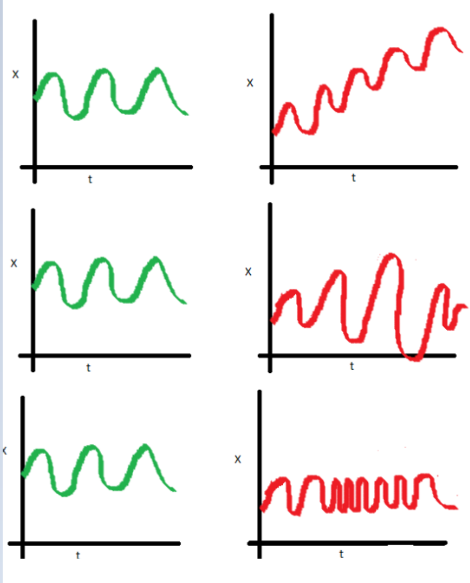
3.1 Stationarity

But in order to be able to use VAR techniques we need to be sure that our variables are stationary.

Whith Stationarity we mean stable statistic properties:

- Constant Mean: $\mathrm{E}(X_t) = \mu$, where μ is a constant.
- Constant Variance: $\mathrm{Var}(X_t) = \sigma^2$, where σ^2 is a constant.
- Time-Invariant Covariance: $\mathrm{Cov}(X_t, X_{t-j}) = \gamma(j)$, where $\gamma(j)$ is a function of lag j but does not depend on time t.

3.1 Stationarity



4 Data Transformation

As we can see our variables are currently presenting a none stationary behavior.

So in order to make our variables stationary we are going to do some transformation.

1. First, we will take logarithms of our variables.
2. We will transform our variables into differences.

4.1 Data in Logarithms

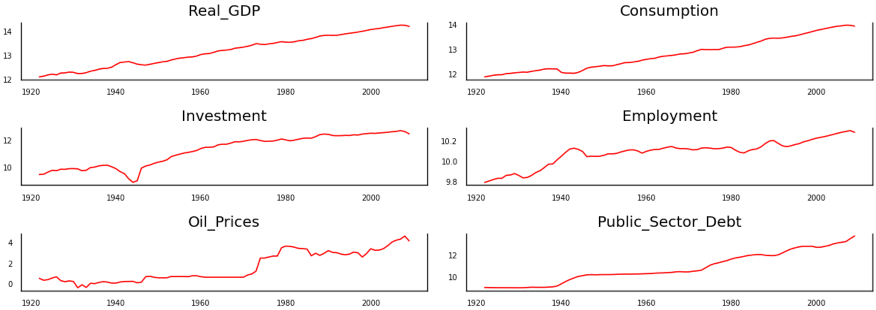


Figure: Figure

4.2 Data in Log-differences

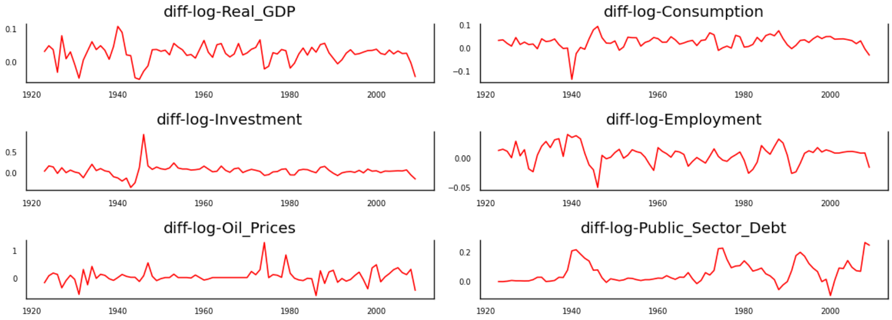


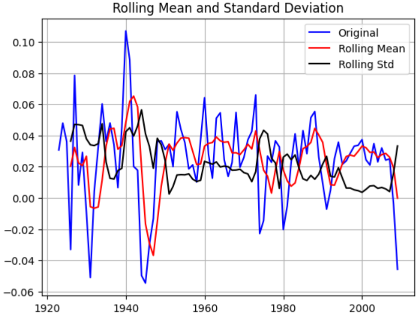
Figure: Figure

- Null Hypothesis (H_0): The time series has a unit root, indicating it is non-stationary.
- Alternative Hypothesis (H_1): The time series does not have a unit root, indicating it is stationary.

4.3 Augmented Dickey–Fuller test

```
Real_GDP
Test Statistic            -5.350074
p-value                    0.000004
Lags Used                  1.000000
Observations Used         85.000000
Critical Value (1%)       -3.509736
Critical Value (5%)       -2.896195
Critical Value (10%)      -2.585258
dtype: float64
```

```
Consumption
Test Statistic            -5.294196
p-value                    0.000006
Lags Used                  0.000000
Observations Used         86.000000
Critical Value (1%)       -3.508783
Critical Value (5%)       -2.895784
Critical Value (10%)      -2.585038
dtype: float64
```



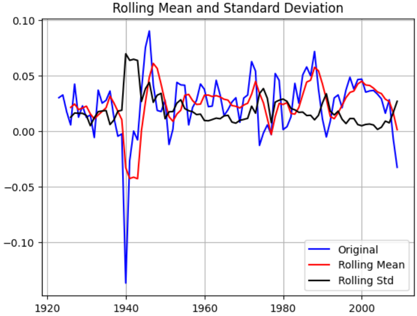


Figure: Figure

Now that we now that our variables are stationary we are in a position of estimating our model

We split our data in two.

- Data Train: include information for the period 1922-2009
- Data Test: the period we want to predict with our model, 2010-2016

.

5.1 Determine the number of lags of our variables

VAR Order	AIC	BIC	FPE
0	-60.30	-59.72	6.516e-27
1	**-63.28***	**-59.81***	**3.357e-28***
2	-63.01	-56.65	4.881e-28
3	-63.27	-54.01	4.990e-28

5.2 The model VAR(1)

Our VAR model is gonna be defined by 10 simultaneous equations, each with one of the variables as the endogenous variables depending on the other variables plus itself with one lag,example:

$$\begin{aligned}
\text{Real_GDP}_t = \alpha_0 + \alpha_1 \text{Real_GDP}_{t-1} + {} & \\
& \beta_1 \text{Consumption}_{t-1} + \beta_2 \text{Investment}_{t-1} + \\
& \beta_3 \text{Employment}_{t-1} + \beta_4 \text{Oil_Prices}_{t-1} + \\
& \beta_5 \text{Public_Sector_Debt}_{t-1} + \beta_6 \text{TOT}_{t-1} + \\
& \beta_7 \text{Total_Managed_Expenditure}_{t-1} + \\
& \beta_8 \text{M1}_{t-1} + \beta_9 \text{HousePrice}_{t-1} + \epsilon_t
\end{aligned}$$

5.3 VAR Results

Durbin-Watson

After getting the results we run a D-W correlation residual test within the residuals of the equations to see if there is any pattern not captured by our equations, and we confirmed is not the case.

Accuracy

Then we evaluated our results using the Mean absolute percentage error to measure of prediction accuracy of our model and we got an 0.883.

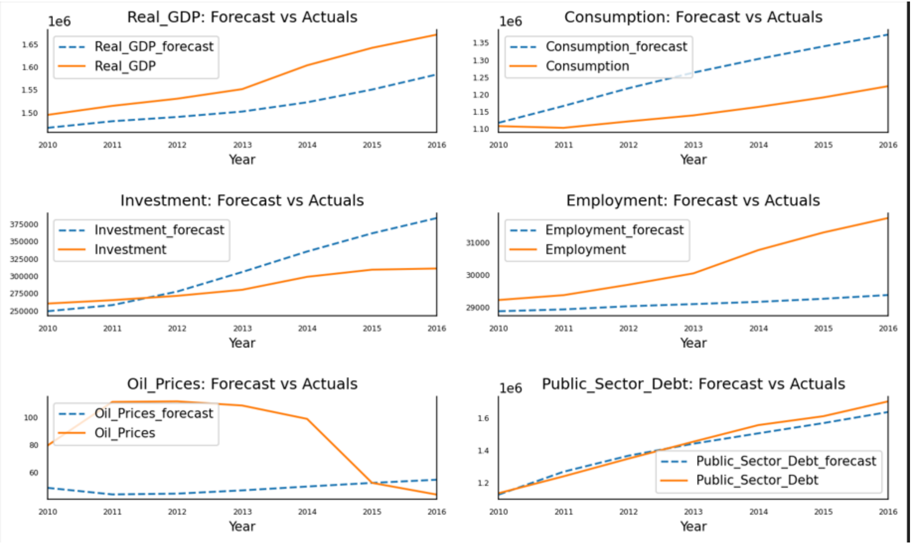


Figure: Figure

6 VARMAX model

Following the same logic we are going to estimate a similar model call VARMAX

This model is essentially similar to the previous one, but now we will add 2 elements:

1. A Moving Average component to try to capture a past behavior of some shocks.
2. A trend, that can be null, constant, linear or both

6.1 Determine the number of lags and elements of our model

p	q	tr	AIC
1	1	c	-2.427211e+03
2	1	t	-2.412391e+03
0	1	n	-2.398573e+03
15	2	ct	-2.343128e+03
12	2	n	-2.320649e+03

6.2 The model VARMAX(1,1,c)

Our VARMAX model is gonna be defined by 10 simultaneous equations, each one with one of the variables as the endogenous variables depending on the other variables with one lag plus a MA(1) component and a constant.

$$\begin{aligned}\text{Real\_GDP}_t = \alpha_0 + \alpha_1 \text{Real\_GDP}_{t-1} + \\ \beta_0 + \beta_1 \text{Consumption}_{t-1} + \beta_2 \text{Investment}_{t-1} + \\ \beta_3 \text{Employment}_{t-1} + \beta_4 \text{Oil\_Prices}_{t-1} + \\ \beta_5 \text{Public\_Sector\_Debt}_{t-1} + \beta_6 \text{TOT}_{t-1} + \\ \beta_7 \text{Total\_Managed\_Expenditure}_{t-1} + \\ \beta_8 \text{M1}_{t-1} + \beta_9 \text{HousePrice}_{t-1} + \theta_{\mathbf{1}} \epsilon_{\mathbf{t-1}} + \epsilon_t\end{aligned}$$

6.3 VARMAX Results

Durbin-Watson

After getting the results we run a D-W correlation residual test within the residuals of the equations to see if there is any pattern not captured by our equations, and we confirmed is not the case.

Accuracy

The Mean absolute percentage error equal to 0.887

6.3 VARMAX Results

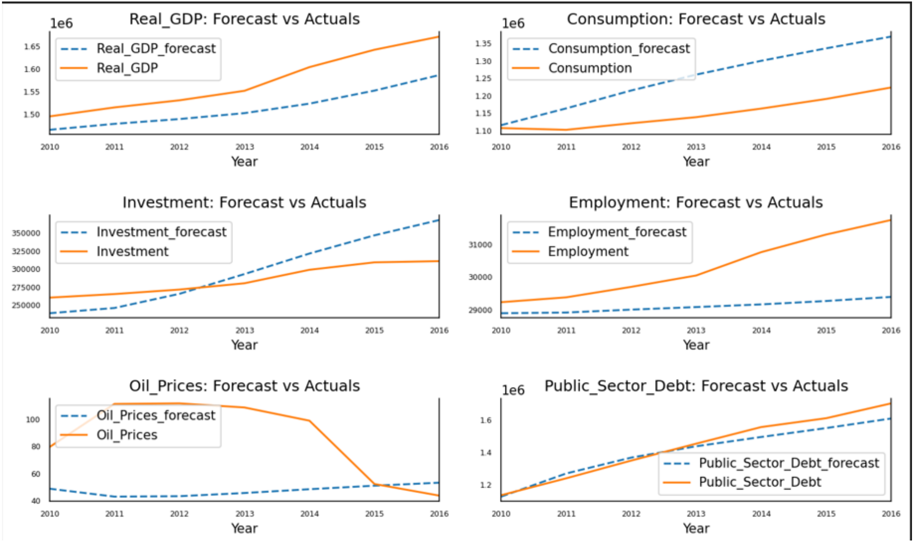


Figure: Figure

7 VECM model

Finally we are going to estimate Vector of error correction model.

This models use stationary variables that are co-integrated between themselves i.e exist a long run relation between themselves that produce stationary results.

In this case we are not going to use our variables in differences since the idea of the model is to find the linear combinations of parameters that create stationary results.

Figure: Figure

7.2 VECM Results

Durbin-Watson

After getting the results we run a D-W correlation residual test within the residuals of the equations to see if there is any pattern not captured by our equations, and we confirmed is not the case.

Accuracy

The Mean absolute percentage error equal to 0.869

7.2 VECM Results

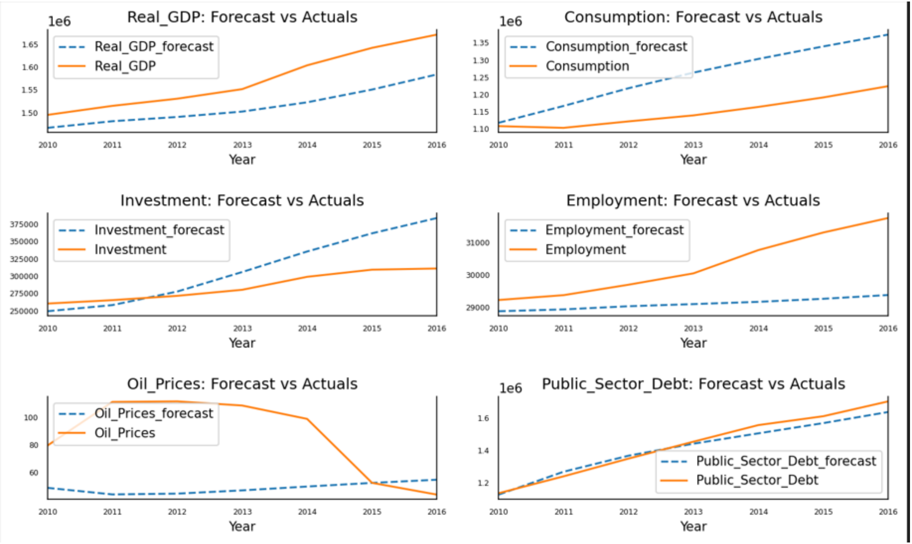


Figure: Figure

8 Conclusions

The main interest of VAR modellings is to look at the dynamics effects between different variables and being able to predict their behavior simultaneously.
After comparing the results of the 3 VAR models we have determined that the VARMAX is the one who capture the best the behavior of the economy.

Our models are more pessimistic than reality regarding the GDP and employment and produce poor predictions on more volatile variables like Oil prices.

Still further improvements can be made by choosing other macroeconomic variables or by reviewing the macroeconometric literature regarding time series prediction.

Also it would be nice to compare the results that VAR provided with other methodologies such ass Artificial Neural Network (ANN) Models or State

¡Thank you!