VAR Models for Macroeconomic Forecasting

VAR models are interesting econometric tool for explaining how different time series are influenced by each other, and are a very useful tool for econometric forecasting. They can often help answer the question: given a sudden change in one time series, how might others respond, given how they have behaved with each other in the past?

Let's start as always by importing some packages we'll need along the way:

```
require( dplyr )
require( lubridate )
require( fImport )
require( vars )
```

Gathering the Data

The first step here is to collect the historical macroeconomic data that we'll need.

```
CPI.Data      = read.csv( 'CPIAUCSL.csv' , header = TRUE )
FEDFUNDS.Data = read.csv( 'FEDFUNDS.csv' , header = TRUE )
UNRATE.Data   = read.csv( 'UNRATE.csv'   , header = TRUE )
```

Import the data from CSVs that we downloaded from the FRED webpage, and then join them together. To limit our time range we also extract only the window since January 2007.

```
Joined.Data =
  CPI.Data %>%
  left_join( FEDFUNDS.Data , by = 'DATE'  ) %>%
  left_join( UNRATE.Data   , by = 'DATE' ) %>%
  na.omit %>%
  mutate( Format_Date = as.Date( as.character( DATE ) ) ) %>%
  dplyr::filter( Format_Date >= as.Date( '1965-01-01' ) ) %>%
  dplyr::select( -one_of( c( 'DATE' , 'Format_Date'  ) ) )

head( Joined.Data )
```

```
##   CPIAUCSL FEDFUNDS UNRATE
## 1    31.28     3.90    4.9
## 2    31.28     3.98    5.1
## 3    31.31     4.04    4.7
## 4    31.38     4.09    4.8
## 5    31.48     4.10    4.6
## 6    31.61     4.04    4.6
```

Fitting the VAR Model

Next we can use the VARselect() function from the vars package, which will help us identify the optimal lag according to a number of different criterion:

```
VAR.Leg.Recommendations <- VARselect( Joined.Data , lag.max=12, type="const")

VAR.Leg.Recommendations$selection
```

```
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      9      4      3      9
```

For this exercise we'll choose to fit the VAR according to the third option, where SC stands for the Scwarz Criterion. And so you can see in the next chunk that when we fit the VAR model, we set the numbers of lags, p, to be equal to the numbers of lags recommended by SC above:

```
VAR.Fitted.Model <- VAR( Joined.Data, p = VAR.Leg.Recommendations$selection[3] , type="const" )

options( show.signif.stars = FALSE )

summary( VAR.Fitted.Model )
```

```
##
## VAR Estimation Results:
## =========================
## Endogenous variables: CPIAUCSL, FEDFUNDS, UNRATE
## Deterministic variables: const
## Sample size: 637
## Log Likelihood: -405.261
## Roots of the characteristic polynomial:
##     1 0.9845 0.9664 0.5852 0.4262 0.4262 0.4159 0.4159 0.4064
## Call:
## VAR(y = Joined.Data, p = VAR.Leg.Recommendations$selection[3],
##     type = "const")
##
##
## Estimation results for equation CPIAUCSL:
## =========================================
## CPIAUCSL = CPIAUCSL.l1 + FEDFUNDS.l1 + UNRATE.l1 + CPIAUCSL.l2 + FEDFUNDS.l2 + UNRATE.l2 + CPIAUCSL.l
##
##             Estimate Std. Error t value Pr(>|t|)
## CPIAUCSL.l1  1.49024    0.03943  37.799  < 2e-16
## FEDFUNDS.l1  0.07219    0.03066   2.354   0.0189
## UNRATE.l1    0.04764    0.08726   0.546   0.5853
## CPIAUCSL.l2 -0.66565    0.06643 -10.020  < 2e-16
## FEDFUNDS.l2 -0.03234    0.04945  -0.654   0.5133
## UNRATE.l2   -0.08827    0.12998  -0.679   0.4973
## CPIAUCSL.l3  0.17660    0.03955   4.465 9.48e-06
## FEDFUNDS.l3 -0.01503    0.03000  -0.501   0.6165
## UNRATE.l3    0.04234    0.08857   0.478   0.6328
## const       -0.06526    0.07423  -0.879   0.3796
##
##
## Residual standard error: 0.3631 on 627 degrees of freedom
## Multiple R-Squared:     1,   Adjusted R-squared:     1
## F-statistic: 2.533e+06 on 9 and 627 DF,  p-value: < 2.2e-16
##
##
## Estimation results for equation FEDFUNDS:
## =========================================
## FEDFUNDS = CPIAUCSL.l1 + FEDFUNDS.l1 + UNRATE.l1 + CPIAUCSL.l2 + FEDFUNDS.l2 + UNRATE.l2 + CPIAUCSL.l
##
##             Estimate Std. Error t value Pr(>|t|)
## CPIAUCSL.l1  0.04643    0.05116   0.908 0.364490
```

```
## FEDFUNDS.l1  1.41389    0.03979  35.535  < 2e-16
## UNRATE.l1   -0.61405    0.11324  -5.422 8.40e-08
## CPIAUCSL.l2 -0.06259    0.08621  -0.726 0.468098
## FEDFUNDS.l2 -0.61124    0.06417  -9.525  < 2e-16
## UNRATE.l2    0.57589    0.16867   3.414 0.000681
## CPIAUCSL.l3  0.01532    0.05132   0.298 0.765480
## FEDFUNDS.l3  0.18004    0.03893   4.624 4.57e-06
## UNRATE.l3    0.02785    0.11494   0.242 0.808642
## const        0.25201    0.09633   2.616 0.009107
##
##
## Residual standard error: 0.4713 on 627 degrees of freedom
## Multiple R-Squared: 0.9848,  Adjusted R-squared: 0.9846
## F-statistic:  4507 on 9 and 627 DF,  p-value: < 2.2e-16
##
##
## Estimation results for equation UNRATE:
## =======================================
## UNRATE = CPIAUCSL.l1 + FEDFUNDS.l1 + UNRATE.l1 + CPIAUCSL.l2 + FEDFUNDS.l2 + UNRATE.l2 + CPIAUCSL.l3
##
##              Estimate Std. Error t value Pr(>|t|)
## CPIAUCSL.l1 -0.012017   0.017711  -0.678   0.4977
## FEDFUNDS.l1  0.001800   0.013774   0.131   0.8961
## UNRATE.l1    1.100149   0.039201  28.064  < 2e-16
## CPIAUCSL.l2 -0.001683   0.029843  -0.056   0.9551
## FEDFUNDS.l2 -0.012745   0.022214  -0.574   0.5664
## UNRATE.l2    0.105740   0.058390   1.811   0.0706
## CPIAUCSL.l3  0.013984   0.017767   0.787   0.4315
## FEDFUNDS.l3  0.020953   0.013478   1.555   0.1205
## UNRATE.l3   -0.215641   0.039791  -5.419 8.54e-08
## const       -0.023708   0.033346  -0.711   0.4774
##
##
## Residual standard error: 0.1631 on 627 degrees of freedom
## Multiple R-Squared: 0.9903,  Adjusted R-squared: 0.9902
## F-statistic:  7142 on 9 and 627 DF,  p-value: < 2.2e-16
##
##
##
## Covariance matrix of residuals:
##           CPIAUCSL FEDFUNDS    UNRATE
## CPIAUCSL  0.131867  0.01081 -0.002845
## FEDFUNDS  0.010814  0.22208 -0.013490
## UNRATE   -0.002845 -0.01349  0.026612
##
## Correlation matrix of residuals:
##          CPIAUCSL FEDFUNDS   UNRATE
## CPIAUCSL  1.00000  0.06319 -0.04802
## FEDFUNDS  0.06319  1.00000 -0.17548
## UNRATE   -0.04802 -0.17548  1.00000
```

Notice the equations given for each variable in the summary output, because this is one of the elements that I like about the VAR model for explaining and forecasting time series: each variable is a function of both itself and all other variables in the system over the last several time periods. Intuitively, this makes sense and helps us tell a story:

Said in words, our VAR model is saying that the unemployment rate this month is both a function of all other variables last month AND the month before that AND the month before that. The same can be said for any of the other variables in the system.

Impulse Response Functions

The IRF or impulse response function allows us to see the change in the system from an impulse response (aka a "shock") in a given variable. So we shock one variable and can watch how the others respond.

```
months.ahead = 12

UNRATE.Shock.Object = irf( VAR.Fitted.Model , impusle = "UNRATE" , n.ahead = months.ahead  )

UNRATE.Shock.Object$irf$UNRATE
```

```
##           CPIAUCSL   FEDFUNDS    UNRATE
##  [1,]  0.000000000  0.0000000 0.1604885
##  [2,]  0.007645083 -0.0985481 0.1765613
##  [3,] -0.001476875 -0.1549745 0.2109444
##  [4,] -0.014032298 -0.1828081 0.2171138
##  [5,] -0.026084348 -0.1888491 0.2229481
##  [6,] -0.034796922 -0.1895221 0.2218046
##  [7,] -0.041718724 -0.1874009 0.2192751
##  [8,] -0.047869652 -0.1839815 0.2149284
##  [9,] -0.053811054 -0.1793699 0.2100543
## [10,] -0.059558403 -0.1741354 0.2047719
## [11,] -0.065082828 -0.1686111 0.1993978
## [12,] -0.070375032 -0.1630352 0.1940109
## [13,] -0.075460102 -0.1575080 0.1886971
```

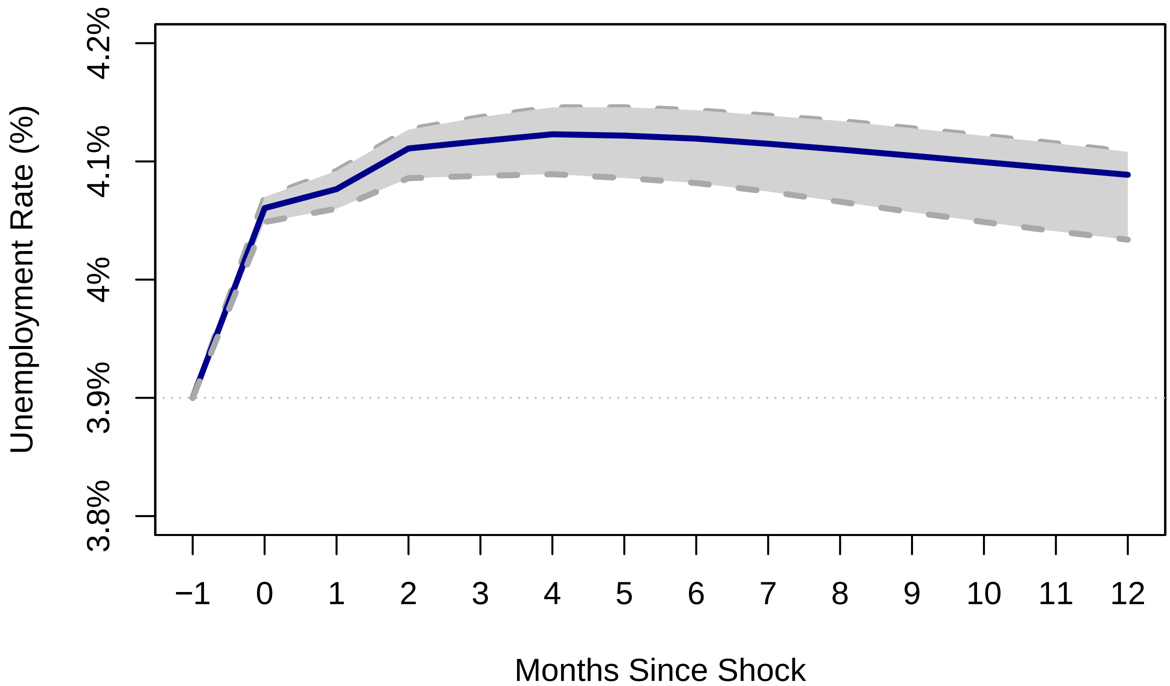
The data contained in the IRF object is in relative terms, meaning that each of those numbers is an addition to its last value. So we can make some casual observations from this output – as UNRATE is shocked upwards or increases (all the numbers in the UNRATE column are positive), then the other two time series typically lower (the changes in the FEDFUNDS and CPI columns are both mostly negative).

We can also do some conversion of these relatives changes into absolute levels, in order to make this a more useful forecasting tool:

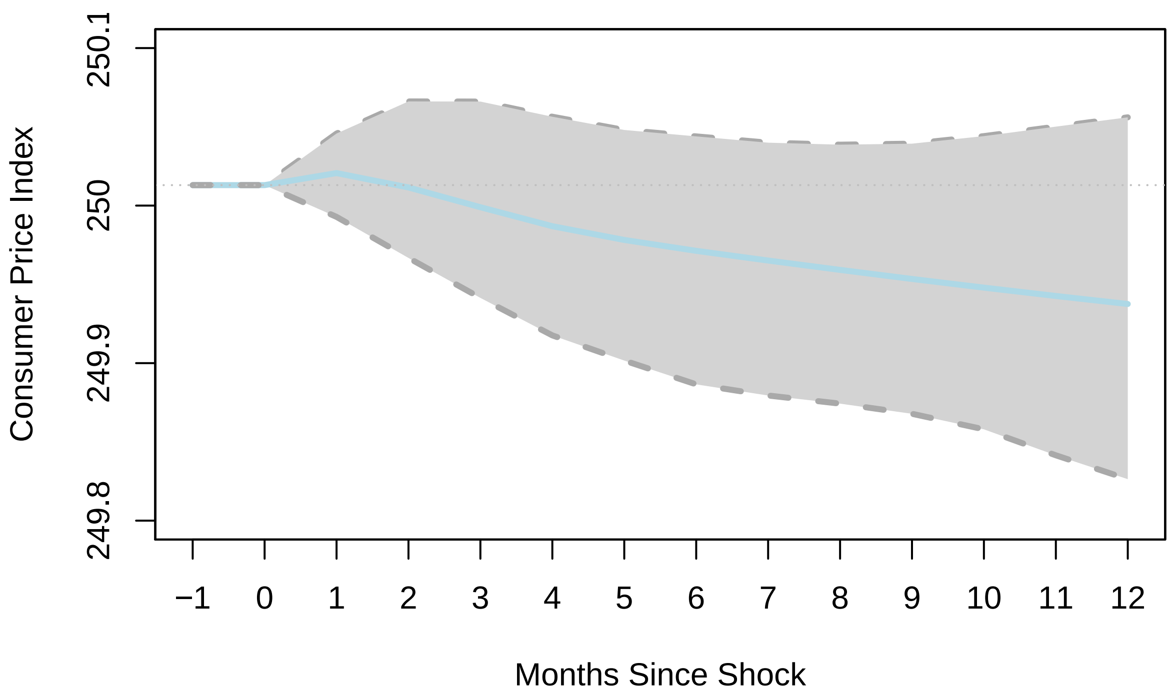
```
Upper.IRF.Levels =
  tail( Joined.Data , 1 ) %>%
  rbind( UNRATE.Shock.Object$Upper$UNRATE ) %>%
  lapply( FUN = function(x){ c( x[1] , x[2:length(x)] + x[1] )  } )

Middle.IRF.Levels =
  tail( Joined.Data , 1 ) %>%
  rbind( UNRATE.Shock.Object$irf$UNRATE ) %>%
  lapply( FUN = function(x){ c( x[1] , x[2:length(x)] + x[1] )  } )

Lower.IRF.Levels =
  tail( Joined.Data , 1 ) %>%
  rbind( UNRATE.Shock.Object$Lower$UNRATE ) %>%
  lapply( FUN = function(x){ c( x[1] , x[2:length(x)] + x[1] )  } )
```

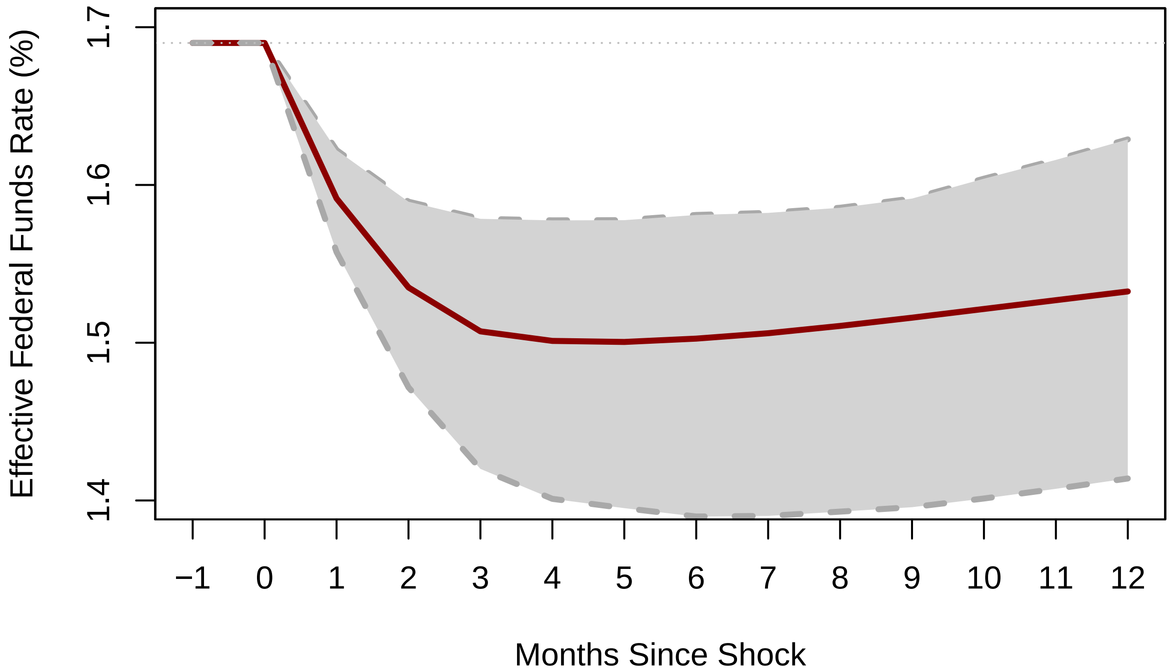
With that done we can go ahead and plot the results – let's do it first for our shocked variable, the Unemployment Rate:

```
x.sequence = 0:(months.ahead+1) - 1

ylim = c( 3.8 , 4.2 )

plot( y = Upper.IRF.Levels$UNRATE
      , x = x.sequence
      , col = 'darkgray'
      , type = 'l', lty = 2, lwd = 3
      , xlab = NA , xaxt = 'n'
      , ylim = ylim, ylab = NA , yaxt = 'n'
)

polygon( x = c( x.sequence , rev( x.sequence ) )
         , y = c( Upper.IRF.Levels$UNRATE , rev( Lower.IRF.Levels$UNRATE ) )
         , border = NA
         , col = 'lightgrey'
)

par( new = TRUE )

plot( y = Middle.IRF.Levels$UNRATE
      , x = x.sequence
      , col = 'darkblue'
      , type = 'l', lty = 1, lwd = 3
      , xlab = NA , xaxt = 'n'
      , ylim = ylim, ylab = NA , yaxt = 'n'
)

par( new= TRUE )

plot( y = Lower.IRF.Levels$UNRATE
      , x = x.sequence
      , col = 'darkgray'
      , type = 'l', lty = 2, lwd = 3
      , xlab = NA , xaxt = 'n'
      , ylim = ylim, ylab = NA , yaxt = 'n'
)

abline( h = head( Middle.IRF.Levels$UNRATE , 1 ) , lwd = 1 , lty = 3 , col = 'gray' )

mtext( text = 'Unemployment Rate (%)' , side = 2 , line = 3 )
y.marks = seq( from = ylim[1] , to = ylim[2] , by = 0.1 )
axis( side = 2 , at = y.marks , labels = paste0( y.marks , '%') )

mtext( text = 'Months Since Shock' , side = 1 , line = 3 )
axis( side = 1 , at = x.sequence , labels = x.sequence )
```



Now let's do it for the other variables:





Here we see a response close to what most might expect, in that as unemployment rises:

- The Federal Funds rate declines, as the FED lowers rates to stimulate growth.
- The Consumer Price Index (CPI) decreases. Lower employment results in less pressure to increase consumer prices.

Conclusion

Here we've demonstrated how to use a VAR model in order to conduct a sort of macroecnomic sensitivity analysis. This type of model is useful because of its practicality (it is calibrated based on historical data and is therefore representative of real observed relationship), and also because of its versatility (we used three variables here, but many more time series can be added to the system, depending on the sensitivities you are interested in testing).