Assignment 3 Solution

1 Analytical Exercise

1. Let $\mathrm{y}_t = (X_t, Y_t)'$ be given by

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

$$Y_t = h_1 X_{t-1} + u_t$$

where $(\varepsilon_t, u_t)'$ is a vector white noise with contemporaneous variance-covariance matrix given by

$$\begin{bmatrix} E(\varepsilon_t^2) & E(\varepsilon_t u_t) \\ E(u_t\varepsilon_t) & E(u_t^2) \end{bmatrix} = \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}$$

Calculate the autocovariance matrices $\{\Gamma_k\}_{k=-\infty}^{\infty}$ for this process.

Answer:

$$y_t = (X_t, Y_t)'$$

$$X_t = \varepsilon_t + \theta\varepsilon_{t-1}$$

$$Y_t = h_1 X_{t-1} + u_t$$

$$\Rightarrow Y_t = h_1\varepsilon_{t-1} + h_1\theta\varepsilon_{t-2} + u_t$$

The variances are calculated as

$$\Gamma_0 = E\left[(y_t - \mu)(y_t - \mu)'\right]$$

$$\Gamma_0 = E\left[\begin{array}{c} (\varepsilon_t + \theta\varepsilon_{t-1}) \\ (h_1\varepsilon_{t-1} + h_1\theta\varepsilon_{t-2} + u_t) \end{array} \quad (\varepsilon_t + \theta\varepsilon_{t-1}) \quad (h_1\varepsilon_{t-1} + h_1\theta\varepsilon_{t-2} + u_t) \right]$$

$$\Gamma_0 = \begin{bmatrix} (\sigma_\varepsilon^2 + \theta^2\sigma_\varepsilon^2) & h_1\theta\sigma_\varepsilon^2 \\ h_1\theta\sigma_\varepsilon^2 & (h_1^2\sigma_\varepsilon^2 + h_1^2\theta^2\sigma_\varepsilon^2 + \sigma_u^2) \end{bmatrix}$$

The autocovariances are calculated as

$$\Gamma_1 = E\left[(y_t - \mu)(y_{t-1} - \mu)^{'}\right]$$

$$\Gamma_1 = E\left[\begin{array}{c} (\varepsilon_t + \theta\varepsilon_{t-1}) \\ (h_1\varepsilon_{t-1} + h_1\theta\varepsilon_{t-2} + u_t) \end{array} \quad (\varepsilon_{t-1} + \theta\varepsilon_{t-2}) \quad (h_1\varepsilon_{t-2} + h_1\theta\varepsilon_{t-3} + u_{t-1})\right]$$

$$\Gamma_1 = \begin{bmatrix} \theta\sigma_\varepsilon^2 & 0 \\ h_1\sigma_\varepsilon^2 + h_1\theta\sigma_\varepsilon^2 & h_1^2\theta\sigma_\varepsilon^2 \end{bmatrix}$$

Using the same argument

$$\Gamma_2 = E\left[(y_t - \mu)(y_{t-2} - \mu)^{'}\right]$$

$$\Gamma_2 = \begin{bmatrix} 0 & 0 \\ h_1\theta\sigma_\varepsilon^2 & 0 \end{bmatrix}$$

$$\Gamma_k = 0, k > 2; \Gamma_{-k} = \Gamma_k^{'} fork \leq 2$$

2. Consider the following bivariate VAR

$$y_{1t} = 0.3y_{1,t-1} + 0.8y_{2,t-1} + \varepsilon_{1t}$$

$$y_{2t} = 0.9y_{1,t-1} + 0.4y_{2,t-1} + \varepsilon_{2t}$$

Is this system covariance stationary?

Answer:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 0.3 & 0.8 \\ 0.9 & 0.4 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}$$

$$\Rightarrow Y_t = FY_{t-1} + v_t$$

The above VAR system is stationary if the eigenvalues of F matrix is less than 1. To determine the eigenvalue of F, we know that if λ is eigenvalue of F, then determinant of the matrix

$$\begin{bmatrix} \lambda - 0.3 & -0.8 \\ -0.9 & \lambda - 0.4 \end{bmatrix} = 0$$
$$\Rightarrow \lambda^2 - 0.7\lambda - 0.6 = 0$$

Solving this quadratic equation, we get $\lambda_1 = 1.2 > 1$ and $\lambda_2 = -0.5$. As one of the eigenvalues is greater than 1, this VAR model is non-stationary.

3. Consider the bivariate VAR(p) model

$$A(L)y_t = \varepsilon_t, \varepsilon_t \tilde{\ } iid(0, \Sigma)$$
$$A(L) = I - A_1 L - A_2 L^2 - .... - A_p L^p$$

with Wold (moving average) representation

$$y_t = \Psi(L)\varepsilon_t$$

where $\Psi(L) = \sum_{k=0}^{\infty} \psi_k L^k$ and $\psi_0 = I_2$.

(a) Find the moving average coefficients ψ_k for a VAR(1) model.

Answer: Since $\Psi(L) = A(L)^{-1}$

$$\Rightarrow \Psi(L)A(L) = I$$

For VAR(1) model, $A(L) = I - A_1 L$ and $\Psi(L) = 1 + \psi_1 L + \psi_2 L^2 + - - - --$

$$\Rightarrow \Psi(L)A(L) = 1 + \psi_1 L + \psi_2 L^2 + - - - - - - - A_1 L - A_1 \psi_1 L^2 - ...... = I$$

Collecting terms, we find $\psi_k = A_1 \psi_{k-1} = A_1^k$ (See class notes).

(b) Show that the moving average coefficients for a VAR(2) model can be found recursively by

$\psi_0 = I_2, \psi_1 = A_1$

and $\psi_k = A_1 \psi_{k-1} + A_2 \psi_{k-2}, k > 1$

Answer: Use the same idea that $\Psi(L)A(L) = I$. For VAR(2) $A(L) = I - A_1 L - A_2 L^2$. We can collect terms and show that $\psi_k = A_1 \psi_{k-1} + A_2 \psi_{k-2}$.

2 Empirical Exercise

1. **VAR Models**

   This problem considers a bivariate VAR model for $Y_t = (\Delta s_t; fp_t)'$, where s_t is the logarithm of the monthly exchange rate between the US and the UK, $fp_t = f_t - s_t = i_t^{US} - i_t^{UK}$ is the forward premium or interest rate differential, and f_t is the natural logarithm of the 30-day forward exchange rate. The file usuk.txt has monthly data over the period February 1976 through June 1996 as analyzed in Zivot (2000).

   - Column 1: year / month

   - Column 2: f_t

   - Column 3: s_t

   - Column 4: fp_t

   (a) Plot the monthly return Δs_t and the forward premium fp_t over the period March 1976 through June 1996.

   Answer: See the plot vardata0 in attached text.

   (b) Plot the lag correlation matrix and comment.

   Answer: The lag correlation matrix graph suggests that the cross-correlation between return and forward premium is very small. The ACF plot suggests that forward premium is much more persistent than the exchange rate return as the ACF for the forward premium dies out much more slowly.

   (c) Determine the lag length of the VAR using the AIC. Use a maximum lag of 4. Estimate the VAR and comment on the fit.

   Answer: All criteria suggest lag length of 1. The estimated VAR model with one lag suggests that the exchange rate return regression model has very small explanatory power. The results suggest that only 2 percent of the variations in exchange rate returns can be explained by lagged exchange rate return and lagged forward premium. However, we find that lagged forward premium is significant at 10 percent significance level. For forward premium, however, we find that the persistence of the variable makes the lagged forward premium highly significant and as a result R-squared is also 0.82 for that equation.

(d) Test for Granger-causality. Does s_t Granger-cause fp_t? Does fp_t Granger-cause s_t? Comment.

Answer: As found in the regression equation of exchange rate return, we reject the null of no Granger causality at 10 percent significance level from fp_t to Δs_t. We don't reject the null of no Granger causality from Δs_t to fp_t.

(e) Plot the impulse response functions and comment on the results.

Answer: Our results from the estimated VAR model indicated dynamic feedback (weak) from fp_t to Δs_t, and it is evident in the impulse response functions. We find that a shock to fp_t reduces Δs_t and the maximum response happens next period. The response is significant at 10%. Since is persistent, we find that its response to its own shock persist for more than 2 years. We don't find much response to a shock to Δs_t.

(f) Perform forecast error variance decomposition for maximum lag of 12. Comment on your results.

Answer: The forecast error variance decomposition results confirm the finding of the VAR estimation. We find that around 2 % of the forecast error variance at 20-step ahead forecast of Δs_t is explained by the shock in fp_t, but the corresponding number is less than 0.2 percent for the shock to Δs_t.

(g) Now change the ordering of the variables and plot the impulse response functions. Do the results change as compared to part (e).

Answer: If we change the ordering of the variables, we don't observe change in IRFs.

2. **Forecasting jobs growth using excess bond premium, GZ spread, loan officer survey and real house price growth**

   In this exercise we will examine the predictability of jobs growth using three credit condition measures and house price growth. The details of these variables have been discussed in the class. You can also refer to Kishor (2018) for details. Here is the information on data files associated with this exercise:

   demp_rt.csv: real-time jobs growth from 1985:Q1 with data vintage starting from 1995:01. The latest vintage data is for 2018:Q1.

   demp_var contains data for jobs growth data as it appears today. Column 2 contains jobs growth data, the third column is the financing condition from loan officer's survey (SLO hereafter), 4th column is real house price growth. Use jobs growth data from this file as the actual value.

   gz_ebpq3.txt has GZ spread and excess bond premium as the first and the second column. The starting point for all data series is 1985:Q1.

   Forecast sample and Estimation sample: we perform the following forecasting experiment with all the VAR models outlined below. Our first forecasts cover the period 1995:Q1-1995:Q4 and would have been prepared in 1995:Q1 using 1995:Q1 vintage data. The estimation sample for the first forecasts is 1985:Q1-1994:Q4. We then move ahead one quarter, re-estimate the VAR model and forecast 1995:Q2-1996:Q1, etc. Our final set of forecasts, for 2016:Q4 2017:Q3, would have been prepared in 2016:Q3. Note that in this case, we are simply following the conventional real-time VAR estimation where we move along the columns of the real-time data base for each iteration. We consider 1-Q ahead through 4-Q ahead forecasts. In addition to these quarterly forecasts, we also examine the average over next four quarters.

   (a) Create four bivariate VAR models: jobs growth and GZ spread; jobs growth and excess bond premium; jobs growth and SLO; and job growth and real house price growth. Perform the forecasting exercise for the sample period described above. Make sure to choose the optimal lag length for each iteration of the model estimation. Compare the results with the forecasts generated from AR(1) model by reporting the results of the ratio of RMSE from the VAR model to that of the RMSE from the AR model.

Answer: The results from the bivariate VAR models are

Model	Q=1	Q=2	Q=3	Q=4	Q=1-4
AR1	1.000	1.000	1.000	1.000	1.000
+GZ Spread	0.929	0.925	1.025	1.115	0.996
+EBP	0.867	0.850	0.894	0.954	0.865
+SLO	0.893	0.869	0.884	0.914	0.862
+DRHPI	1.022	1.018	1.009	0.999	1.005

(b) Explain your results in part (a).

Answer: The table shows the ratio of root mean squared errors in comparison to a benchmark AR(1) model. The ratio of less than unity implies that the model in question has lower RMSE than the benchmark AR (1) model. The results suggest that inclusion of credit supply indicators in a model with non-farm payroll growth leads to reduction in RMSE at for short and medium horizons (Q=1,2). The inclusion of credit conditions leads to reduction in RMSE at all forecasting horizons. If real house price growth is added to the univariate model of jobs growth as a predictor, forecasting performance worsens at all forecasting horizon.

(c) Now create three trivariate models: jobs growth, GZ spread and real house price growth; jobs growth, excess bond premium and real house price growth; and jobs growth, SLO and real house price growth. Perform the same exercise as in part (a) and report the ratio of RMSE as compared to an AR(1) model.

Answer: The results are shown below

Model	Q=1	Q=2	Q=3	Q=4	Q=1-4
AR1	1.000	1.000	1.000	1.000	1.000
+GZ Spread+DRHPI	0.812	0.785	0.847	0.925	0.793
+EBP+DRHPI	0.852	0.818	0.855	0.899	0.819
+SLO+DRHPI	0.888	0.842	0.840	0.854	0.817

(d) Does inclusion of housing price growth model lead to an improvement in forecast of the bivariate VAR models?

Answer: Combining credit condition indicators with real house price growth leads to an improvement in forecasting performance at all forecasting horizons. The improvement is higher at medium and long forecasting horizons (Q=3-4).

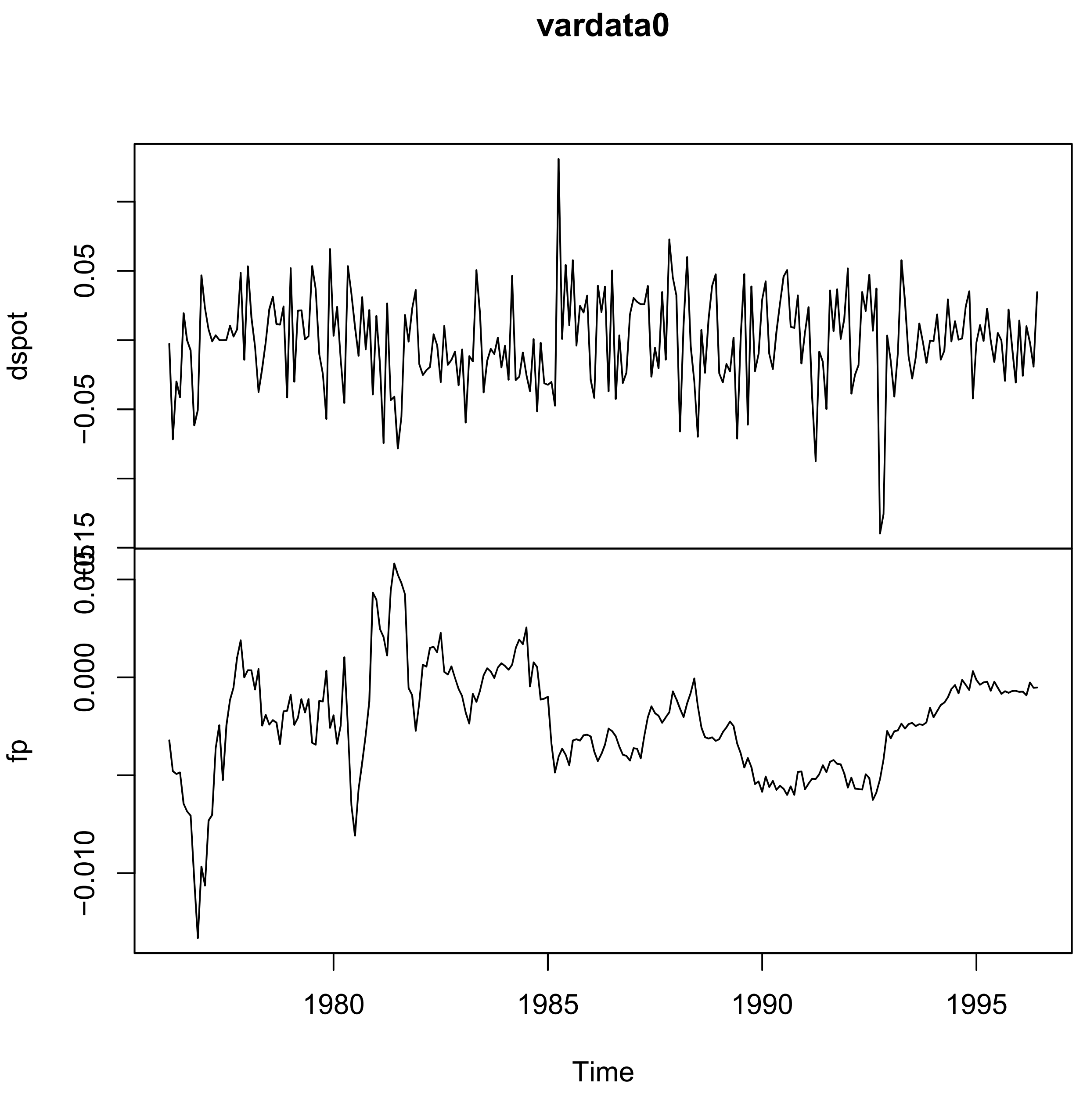
(e) Now combine the forecasts from all the models using simple average, Bates-Granger, Bayesian Model Averaging and Akaike Model Averaging. Report the results for this exercise. Does forecast combination yield superior forecasts? Explain.

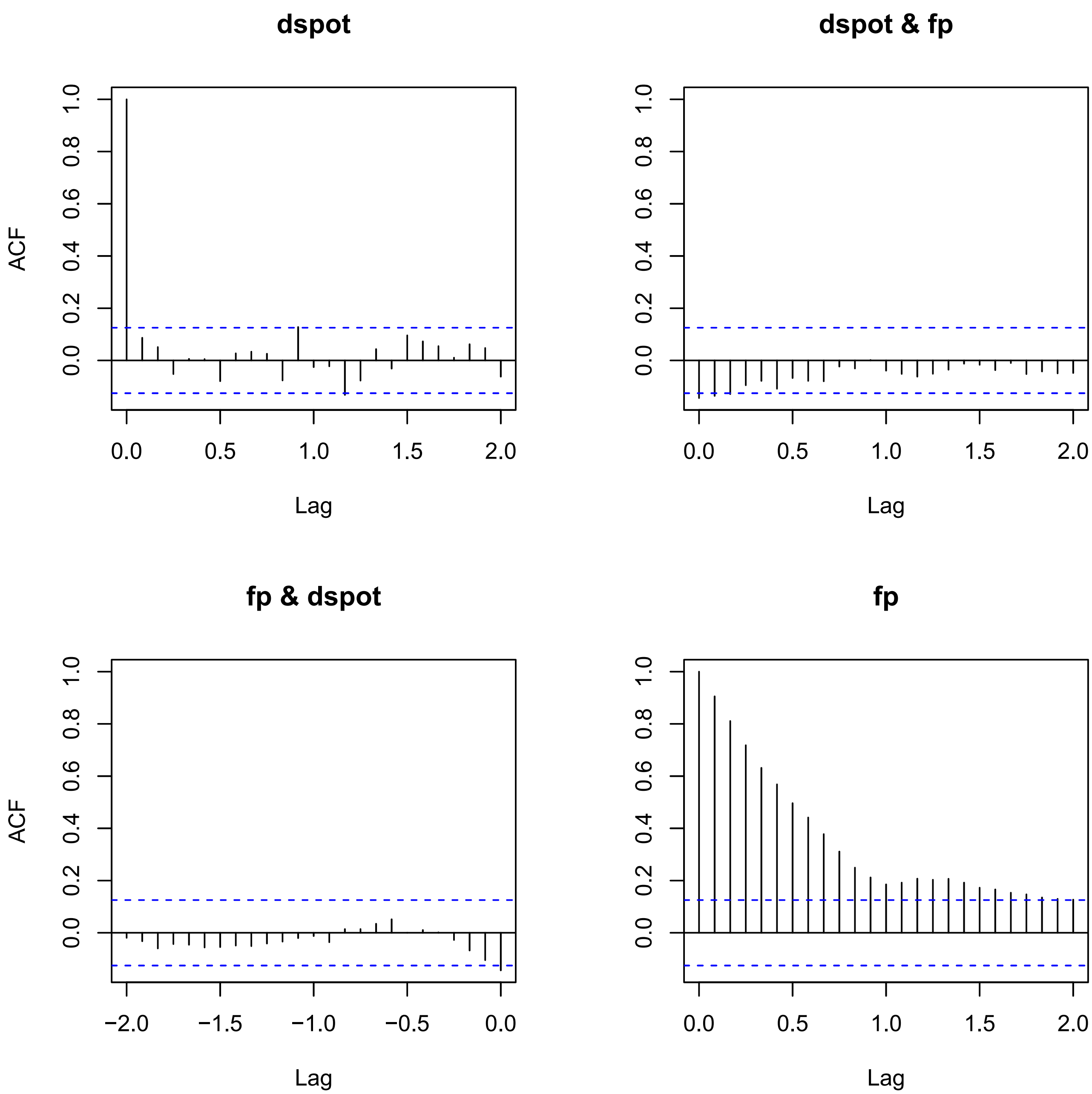
Answer:

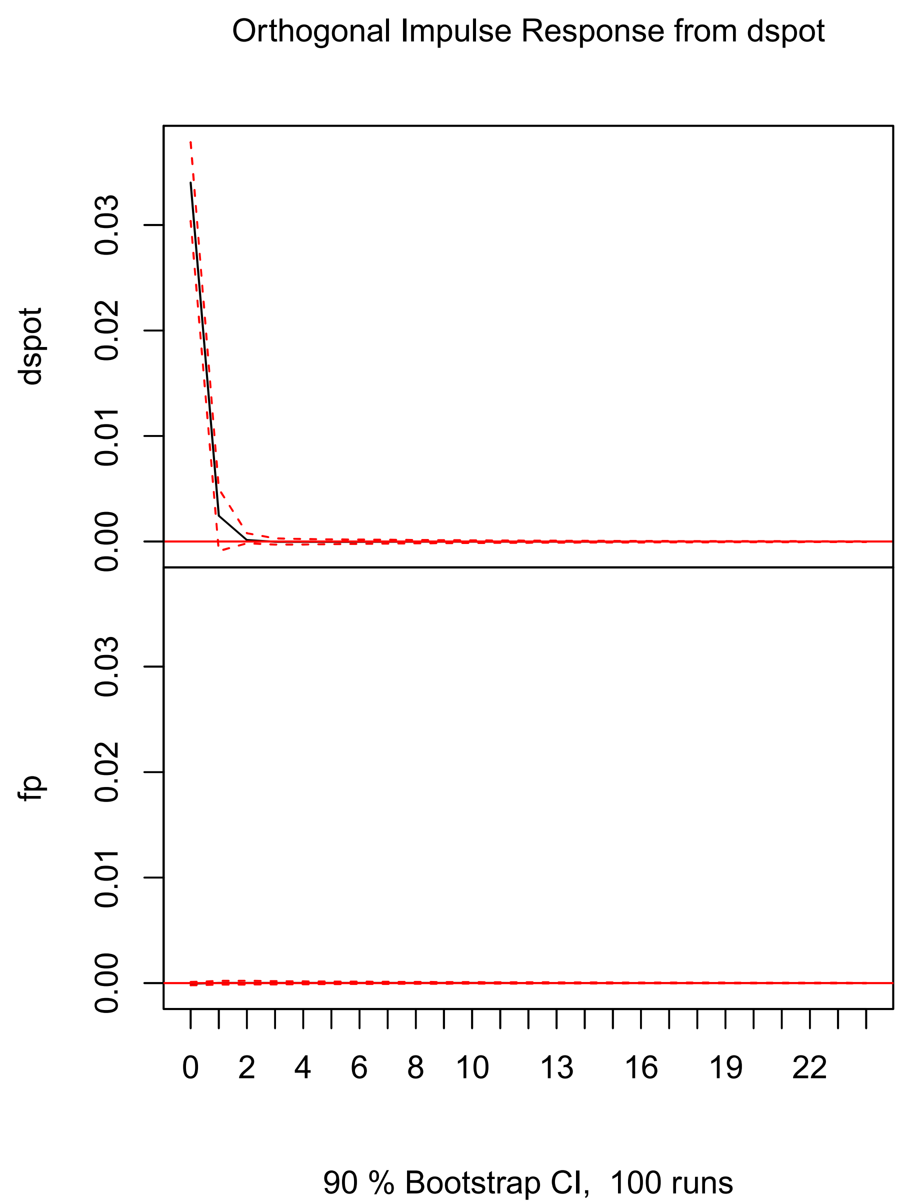
	Q=1	Q=2	Q=3	Q=4	Q=1-4
	Jobs Growth				
AR(1)	1.000	1.000	1.000	1.000	1.000
BG	0.895	0.788	0.811	0.838	0.766
BMA	0.871	0.826	0.856	0.900	0.826
AMA	0.883	0.846	0.882	0.922	0.851
AVERAGE	0.847	0.804	0.836	0.880	0.797

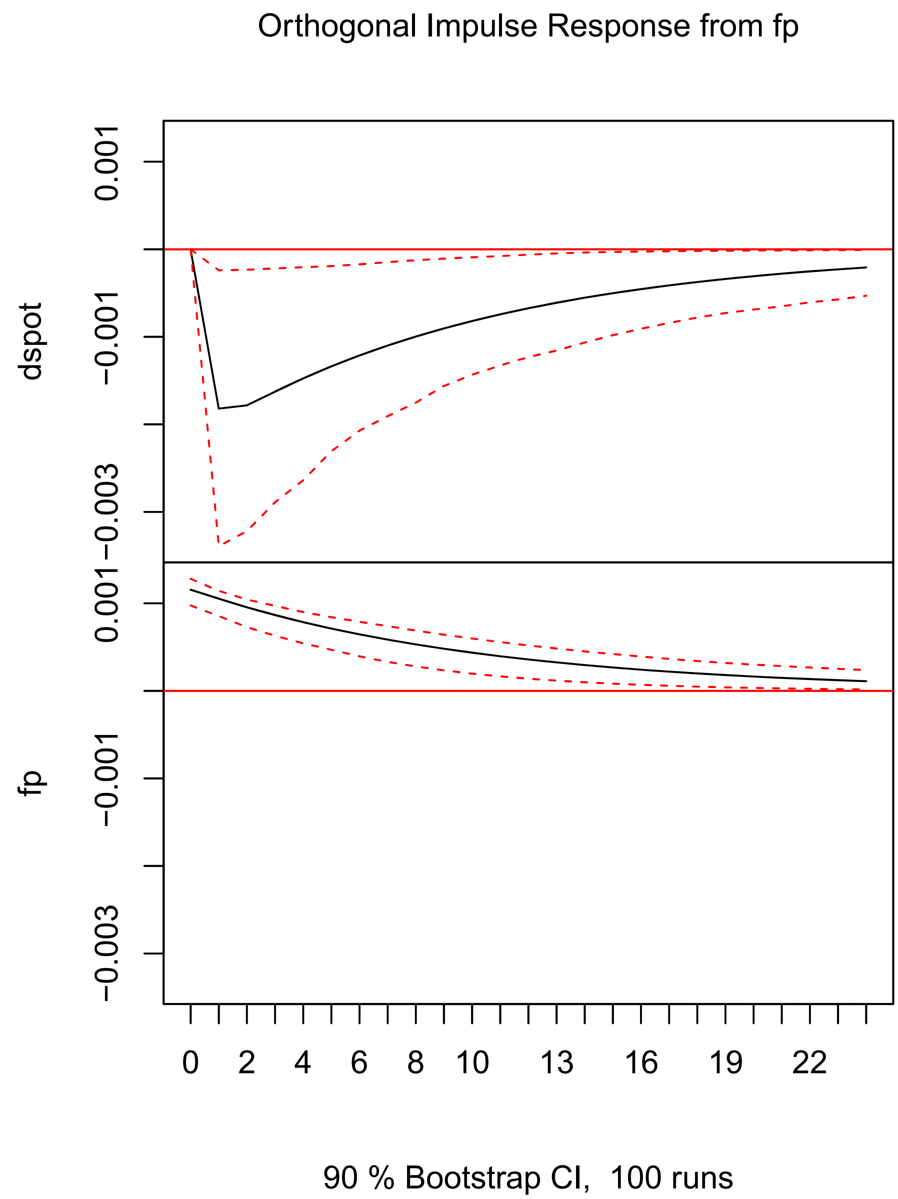
:

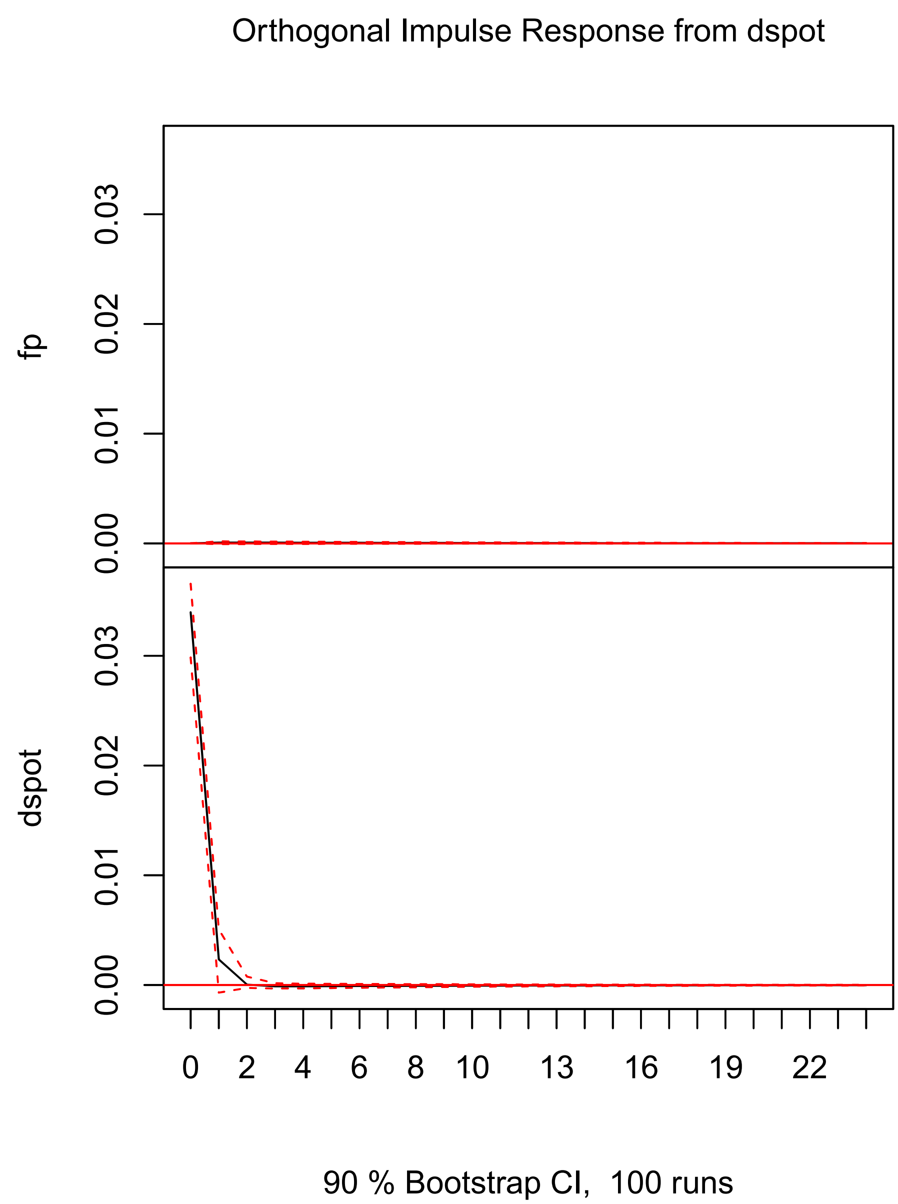
The results for all variables suggest that Bates-Granger (BG) and simple average methods dominate AMA and BMA methods. Especially BG method leads to improvement in forecast over individuals models on average. For example, there is 3% reduction in the ratio of RMSEs from the best VAR model (GZ Spread+DRHPI) for Q=1-4 horizon when BG method is used for forecast combination. Simple averaging also leads to reduction in RMSE for most of the models. One of the reasons why BMA and AMA do not perform as well as BG and simple average is that these two models treat different forecast horizons separately in calculating the combination forecasts, whereas BMA and AMA method uses the same model for different forecast horizons.

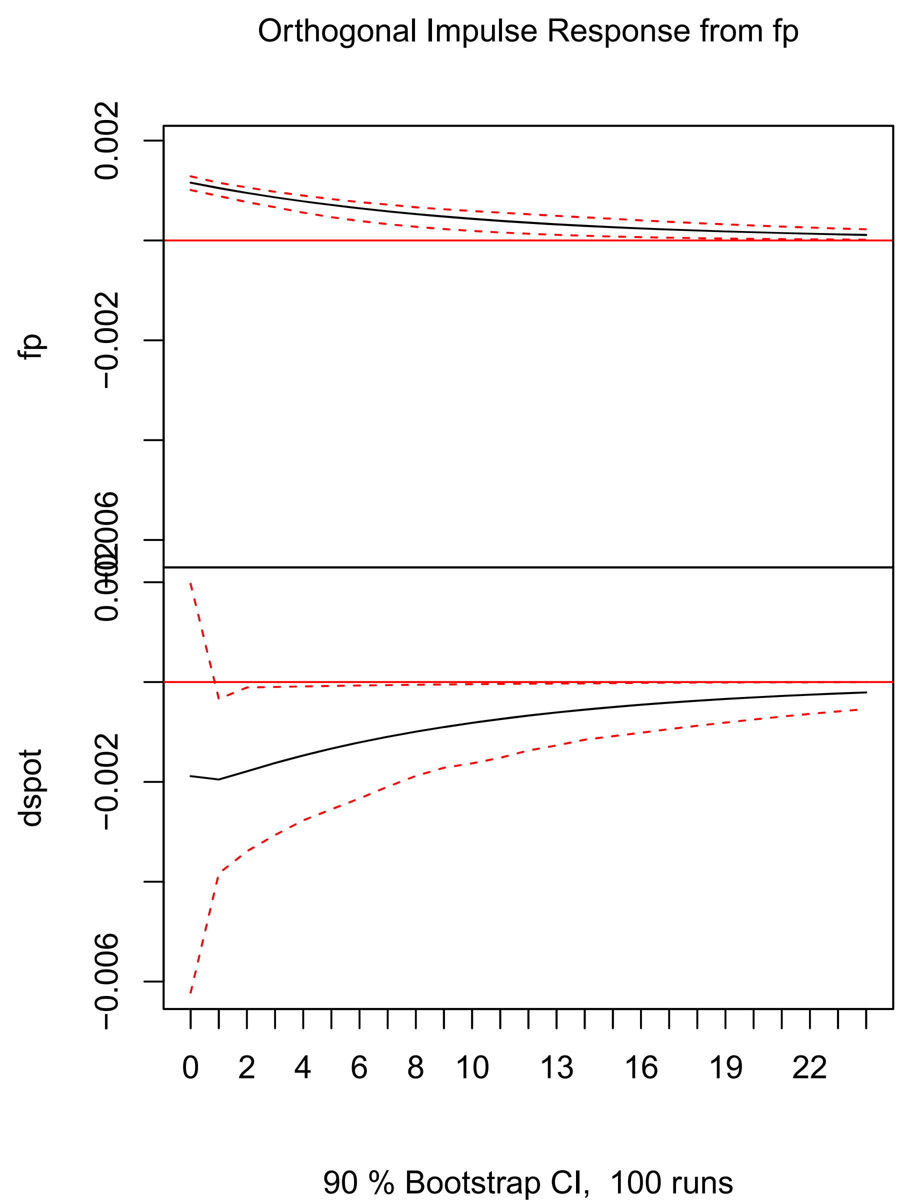

vardata0
dspot
0.05
−0.05
−0.15
fp
0.005
0.000
−0.010
1980
1985
1990
1995
Time


dspot
dspot & fp
fp & dspot
fp
ACF
Lag
0.0
0.2
0.4
0.6
0.8
1.0
−2.0
−1.5
−1.0
−0.5
0.5
1.5
2.0


Orthogonal Impulse Response from dspot
dspot
fp
0.00
0.01
0.02
0.03
0
2
4
6
8
10
13
16
19
22
90 % Bootstrap CI, 100 runs


Orthogonal Impulse Response from fp
dspot
fp
0.001
−0.001
−0.003
0
2
4
6
8
10
13
16
19
22
90 % Bootstrap CI, 100 runs


Orthogonal Impulse Response from dspot
fp
0.00
0.01
0.02
0.03
dspot
0.00
0.01
0.02
0.03
0
2
4
6
8
10
13
16
19
22
90 % Bootstrap CI, 100 runs


Orthogonal Impulse Response from fp
fp
dspot
0.002
−0.002
−0.006
0.002
−0.002
−0.006
0 2 4 6 8 10 13 16 19 22
90 % Bootstrap CI, 100 runs

(1)

```
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
     1      1      1      1

$criteria
                   1             2             3             4
AIC(n) -2.027664e+01 -2.024850e+01 -2.022785e+01 -2.019845e+01
HQ(n)  -2.024158e+01 -2.019007e+01 -2.014604e+01 -2.009326e+01
SC(n)  -2.018963e+01 -2.010348e+01 -2.002481e+01 -1.993740e+01
FPE(n)  1.563027e-09  1.607650e-09  1.641243e-09  1.690272e-09


VAR Estimation Results:
=========================
Endogenous variables: dspot, fp
Deterministic variables: const
Sample size: 243
Log Likelihood: 1778.274
Roots of the characteristic polynomial:
0.9067 0.073
Call:
VAR(y = vardata0, p = 1, type = "const")


Estimation results for equation dspot:
======================================
dspot = dspot.l1 + fp.l1 + const

          Estimate Std. Error t value Pr(>|t|)
dspot.l1  0.068651   0.064653   1.062    0.289
fp.l1    -1.574356   0.809686  -1.944    0.053 .
const    -0.004625   0.002875  -1.609    0.109
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.03402 on 240 degrees of freedom
Multiple R-Squared: 0.02295,    Adjusted R-squared: 0.01481
F-statistic: 2.819 on 2 and 240 DF,  p-value: 0.06165


Estimation results for equation fp:
===================================
```

```
fp = dspot.l1 + fp.l1 + const

           Estimate Std. Error t value Pr(>|t|)
dspot.l1  2.313e-03  2.200e-03   1.051   0.2941
fp.l1     9.110e-01  2.755e-02  33.065   <2e-16 ***
const    -1.901e-04  9.783e-05  -1.943   0.0531 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1


Residual standard error: 0.001158 on 240 degrees of freedom
Multiple R-Squared: 0.822,      Adjusted R-squared: 0.8205
F-statistic: 554.1 on 2 and 240 DF,  p-value: < 2.2e-16



Covariance matrix of residuals:
           dspot         fp
dspot  1.158e-03 -2.182e-06
fp    -2.182e-06  1.340e-06

Correlation matrix of residuals:
         dspot       fp
dspot  1.00000 -0.05539
fp    -0.05539  1.00000



        Granger causality H0: dspot do not Granger-cause fp

data:  VAR object model0
F-Test = 1.1054, df1 = 1, df2 = 480, p-value = 0.2936


        Granger causality H0: fp do not Granger-cause dspot

data:  VAR object model0
F-Test = 3.7807, df1 = 1, df2 = 480, p-value = 0.05243
```